# Educating The Gifted:
## Evaluation Components

Warren D. Wilde
Melvin T. Sillito

# Alberta Education

Alberta
EDUCATION

87-02737/

# EDUCATING THE GIFTED:
# EVALUATION COMPONENTS

Warren D. Wilde
Melvin T. Sillito

Prepared Under Contract to Alberta Education
October 1986

# ABSTRACT

Wilde, W. D. and Sillito, M. T. Educating the Gifted: Evaluation Components, Alberta Education, (April 1986), (83 pp. + app.).

The publication is a guidebook designed for use in evaluating both programs and student growth in the field of education for the gifted. Special problems and concerns are discussed, a strategy is developed for evaluating programs, and procedures are suggested for evaluating student growth. The strategy consists of six sequential procedures: i) developing evaluation questions, ii) identifying appropriate data sources, iii) developing appropriate data gathering procedures, iv) organizing data, v) answering evaluation questions, and vi) reporting answers. Issues and concerns expressed by teachers and found in the literature about evaluating gifted students are addressed. Suggestions for evaluating student achievement in special provisions activities, including peer evaluation and self evaluation techniques, are provided. Some implications for school system evaluation services and other support services are noted.

The guidebook is a product of a project financed mainly by Alberta Education. A consortium of three school systems provided the venue for tryout of the program evaluation strategy and served as a source of information based experience for the suggestions about evaluating student outcomes.

# ACKNOWLEDGEMENTS

The authors are deeply indebted to many. First there are three school boards, Calgary Board of Education, Strathcona County #20, and Camrose School District #1315. These three formed a consortium and contracted with Alberta Education to administer, supervise and assist a project out of which would come four substantive products: a program evaluation and report about the programs for gifted students in each of the three systems, and this guidebook. All four products were to be made available to schools and teachers throughout the province. Strathcona County undertook administration of the contract, and each of the three systems supplied member(s) to the steering committee with Strathcona County providing the chairman. The three school systems assisted greatly in other ways. They provided access to their programs as a venue for program evaluation development and tryout. They facilitated access to schools, administrators, teachers, classrooms, students and parents. They provided relevant documentation and released teachers to assist with solutions to the problems surrounding evaluation of gifted student growth. Obviously the project could not have succeeded without such support and assistance. Nor could it have even been attempted without the fiscal support provided by the Planning Services Branch of Alberta Education. The Planning Services Branch, in addition, provided wise counsel and professional assistance. All of the above administrative personnel gave their assistance willingly, considerately, and with patience. We are most grateful for both the personal and fiscal assistance.

From the above group come five members of the steering committee. Another comes from the Special Education Branch of Alberta Education, and two others from colleagues engaged in concurrent and related projects. The steering committee consisted of: Dr. Grant Jensen, chairman, Assistant Superintendent of Schools, Strathcona County #20; Dr. Clarence Rhodes, Planning Services Branch, Alberta Education; Mr. Allan McLennan, Supervisor of Program Evaluation Services, Calgary Board of Education; Dr. Georgina Adamson, Supervisor of EAS-G, Calgary Board of Education; Mr. William Janzen, Assistant Superintendent, Camrose School District #1315; Dr. Garnet Millar, Special Education Branch, Alberta Education; Dr. Vern Nyberg, Department of Educational Psychology, University of Alberta; Dr. Stanley Clarke, Professor Emeritus, Department of Educational Psychology, University of Alberta. Our thanks in large measure are due to these people for guidance, advice and other help, given in good measure with professional competence and good will.

Another group of people, too numerous for individual mention, made a large and indispensable contribution. The group consists of teachers, program coordinators, administrators and support staff directly involved in educating gifted students. This group included: the Assistant Superintendent and all teachers of gifted students (TOGS) in Camrose; the Assistant Superintendent, Program Coordinator, and all teachers of gifted students in Strathcona County; the entire staff at Oakley Centre, the consultants, curriculum supervisor, itinerant teachers and staff of the Education Assistance Service for the Gifted (EAS-G) component of Calgary's program, and a large sample of school principals and teachers of the gifted involved in special provisions activities for the gifted in the Calgary Board of Education. Their contributions greatly enhance the realism and validity of our efforts. We are very grateful for their help without which the project would have been impossible and the learning meagre.

Finally we express our gratitude to Ruth, Karen, Mary Alice, and others who provided necessary clerical and technical production services as well as to Murray for his assistance in data gathering and his cheerfulness.

W.D.W.

M.T.S.

# PREFACE

The basic premise of education for "the gifted" has pragmatic origins. Hundreds of thousands of teachers teaching millions of children and youth, have noticed a sprinkling who exhibited unusual behavior. The divergent behavior came in a number of patterns. The patterns showed phases of development not clearly related to age but to a degree related to sex. Questioning, curiosity, unique insights and perspectives, even unusual products would often emerge; remain constant in some but submerge in others. For many it became boredom and then chronic somnolence of increasing severity. For others it progressed to frustration and then intolerance, often creating a similar reaction among educators. A large number exhibited only a decline in curiosity. A small, but still uncomfortably large number, discontinued their formal education. A tiny, but twice the expected (and extremely uncomfortable) number took their own lives.

Among the deviants, educators discovered that a significant number had exceptionally high abilities and began to suspect that the other deviants were similarly categorizable. The term gifted, which includes what is commonly referred to as gifted and talented, was adopted as the category designation. When educators began to talk about this group, they naturally decided that something was amiss and needed changing.

The most readily accessible, and blameable, factor for the problem was curriculum. Thus the basic premise in educating gifted children, "they require differentiated curriculum in order to realize their potential contribution to self and to society", was born and has become the operative defining characteristic of schooling for this group.

Important questions about educating gifted students flow from the premise. Some of these questions are:

1. If curricula must be differentiated,
    a) how is it to be done?
       by adding more history? by adding math problems?
       by getting to particle physics sooner?
    b) by whom will it be done?
       home room teachers? by special project teachers?
       by system specialists? by provincial government

specialists?
   - c) in what setting will differentiated curricula be developed and/or tried out? in regular classes? in special classes? in the school? in the school system? in provincial work groups? in demonstration schools?
   - d) who will be involved in deciding curricula? parents? gifted students? teachers? school boards? provincial government?
   - e) who will provide the resources? teachers? schools? Alberta Education?
2. What changes are necessary in teaching/learning activities? What setting is most appropriate?
3. How will goals and objectives be changed? will they be individualized? if so, how? to what extent?
4. Who will decide objectives? parents? gifted students? teachers? the school system? Alberta Education? some of the above? all of them?

Some, mainly pragmatic and expedient, answers have been found for a portion of these questions. Others are receiving varying amounts of attention while some receive none. But none of the questions can be readily dismissed. Pursuit of answers to the above questions and to the ever present pragmatic ones about value of outcomes, especially when the political machinery for raising money is involved, calls for evaluation.

Evaluation of programs and outcomes in provisions for educating gifted students is the focus of this publication. Its point of departure is the basic premise that a differentiated curriculum is necessary. The theme is that the differentiated curriculum necessitates different objectives which in turn demand different data, instruments, and procedures in evaluation of programs and achievement. The destination is a set of suggestions about evaluation.

A number of disclaimers must be added:

1. The guidebook is not another criticism of evaluation practices. It is assumed that for current objectives, current practices and instruments are essentially functional.
2. In similar vein, it is not a critique of current education program evaluation models.
3. It is not a final answer. The careful reader will notice the barely disguised presumption that finality is not now possible, and may not be in the foreseeable future.

With the foregoing in mind, readers engaged in the inspiriting task of working with gifted children may hold hopeful expectation of some useful return for the effort.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................... i

PREFACE .................................. iii

CHAPTER ONE
INTRODUCTION ............................. 1

CHAPTER TWO
EXPECTATIONS ............................. 5

CHAPTER THREE
PROGRAM EVALUATION ....................... 21

CHAPTER FOUR
EVALUATING STUDENT ACHIEVEMENT IN
SPECIAL PROVISIONS ACTIVITIES ............ 51

CHAPTER FIVE
OBSERVATIONS AND IMPLICATIONS
FOR EVALUATION ........................... 76

APPENDIX A
LITERATURE SURVEY
and BIBLIOGRAPHY ......................... 83

APPENDIX B
IPP
(Pilot, 1986, Strathcona County #20) ..... 113

APPENDIX C
FURTHER READING
AND QUICK REFERENCE ...................... 117

APPENDIX D
SAMPLE PLANNING WORKSHEET ................ 120

# CHAPTER ONE

# INTRODUCTION

## ORIGIN OF THE PROJECT

This project is a natural consequence, in the minds of its authors at least, of a previous study, by the same authors, of the state of the current interest and activities in educating gifted students. The previous study drew attention to the need for giving increased attention to the most able and potentially most productive students. No particular claim is being made for perspicacious insight in the previous study. Deficiencies in evaluation practices associated with education programs for gifted students, which formed a basis for the previous study, were all too apparent. If there remained any doubt about the unsatisfactory state of evaluation affairs in education for gifted students, it is quickly dispelled by relevant literature.

The literature indicates that evaluation components range from non-existent to usually inadequate. Many provisions in place are inappropriate to the purposes being served. A majority of program evaluation components are later "add ons" which are, as a consequence, given inadequate planning, resources, and attention. Views on evaluating student related outcomes which are attributable to special educational provisions for gifted students, vary from "unnecessary" to "impossible" to "too time consuming" to "impossible in the short range of the in-school life of students".

Although the need was apparent, how to meet it was much less obvious. Meanwhile pressure to address the need has increased. Alberta Education took steps to encourage provision of appropriate education for gifted students. A task force, advisory to the Minister of Education, met and reported its recommendations. Development of a manual on education of the gifted was undertaken by the Special

Education Branch. Grants were provided which school boards were encouraged to use for teacher in-service and as seed money in establishing special educational provisions for gifted students.

A preliminary exploration with the Planning Services Branch of Alberta Education and with three school systems resulted in a the approval of developing a project proposal to evaluate special provisions for the gifted. The Planning Services Branch agreed to supply the necessary fiscal and other kinds of support. The consortium consisting of Calgary Board of Education, Camrose Public School District, and the County of Strathcona, agreed to supply personnel, classes and consultation. The project began in 1984.

## THE NATURE OF THE PROJECT

The project had three foci. First is evaluation of program provisions, second is evaluation of student outcomes as a consequence of the program provisions, and third is the evaluation component including structure and function which should be put into place in the programs planned.

In each focus the project was designed so as to have both a research and a development component. The research component consisted of a survey of relevant literature and relevant experience, including the experience of the authors and the prior experience of those involved in carrying out the project in the school systems and/or serving on the steering committee. In addition to these was the experience gained during the course of the project itself. The development component included those innovations which could be added by cooperative effort of all involved in working to meet the needs of program and of student outcomes evaluation.

The evaluation of programs and other provisions for educating gifted students was examined in the literature, in light of experience, and through evaluation of the programs in the three school system. These program evaluations were designed to affirm or modify the findings from the literature as well as to address some of the unanswered questions and problems. Concurrent with this study was another one by Dr. V. Nyberg and Dr. S. Clarke (see Appendix A). Findings of this ongoing study in Alberta which had a program evaluation component were provided to the authors of this project. Recourse to acknowledged experts in the field provided still another information resource foundation.

Evaluation of the student outcomes component of this project was similarly based on a literature survey, previous experience, developmental activities, and recourse to experts. In this case, largely because of the complex and infrangible problems, the effects were much less telling. The degree of sophistication discovered was much less complete. Some pivotal help was, however, forthcoming from those offering services in the field.

The project began in April 1984 and continued into the winter of 1985-86. The survey of literature was followed by program evaluation and then by exploration and development of the student evaluation component. It was found necessary to delay the latter component in order to identify those teachers most active in assessing student outcomes.

During the course of the project, consultation with experts in the field provided direction and insight. Dr.Carolyn Callahan from the University of Virginia spent two days in Alberta as special consultant. Dr. Vern Nyberg from the University of Alberta provided ongoing consultation. Dr. Clarence Rhodes of Alberta Education, who served on the steering committee, provided valuable advice and information throughout the course of the project.

## ORGANIZATION OF THE GUIDEBOOK

By intent this guidebook is to be both useful and comprehensive. In preparation for writing the guidebook a lengthy review of the literature was undertaken to serve the reference needs of the authors. This review is included as Appendix A. Only the findings of principle, concept, procedures, constraints, and problems necessary to the guidebook are included therein. These constitute the literature component of the foundation for the guidebook. To this component is added two others: findings from experience, and from expert opinion, most often incorporated without specific reference to the source. These omissions were made in the interests of keeping the publication from becoming an unuseably massive document swollen by the inevitable quotes, paraphrases, and explanations arising from an extensive literature review and detailed reports of field experiences. The three evaluation reports which constitute the major experiential foundation of this guide are published and available from Planning Services Branch, Alberta Education.

The Guidebook itself has five chapters. This first chapter is introductory. Chapter two gives brief consideration to the general problem of norms and expectancies and to the program components, other than those directly related to evaluation, which in education for gifted children differ sufficiently from regular programs to need special attention. These components include: philosophy, definition, identification, curriculum content and delivery, and administration. Chapter three deals with program evaluation. Chapter four with student outcomes evaluation, and chapter five with evaluation provisions as a necessary component of any program for gifted children.

# CHAPTER TWO

# EXPECTATIONS

## ROLE OF EXPECTATIONS

Expectations perform a central, and most often, the controlling role in educational evaluation. Commonly held and broadly understood expectations permeate both the form and substance of public education. Long tradition and massive effort have focussed on curricula which may become quite specific on the concepts and skills teachers expect to teach and which they know parents and school administrators expect them to teach. As students mature they also begin to understand these expectations. Student achievement levels are indices of teacher, school and system effectiveness, so much so that a difference of a few raw score points in means on the standardized measuring instruments is enough to point the finger of inquiry to a teacher and his class, a school, or a school system. The difference is often the difference between complacency and concern as outcomes meet, exceed, or fall short of expected achievement. Departures from the norm in the school day, the school year or the size of classrooms are matters of curiosity. The consequence for evaluation is that the expectations become standards for comparison and assessment. A number of models for evaluation key on these comparisons, giving the category name of discrepancy models. Changes in teaching methods or school organization frequently are assessed by comparing outcomes with normed expectations.

## EXPECTATIONS FOR GIFTED STUDENTS ARE INDEFINITE

In the special programs for gifted children expectations cannot play an equivalent role to those for regular classrooms. There are two main reasons; first, norms for student outcomes expected from the special provisions for gifted students simply do not exist; second, there appears to be a persisting ambivalence in establishing expectancies for the gifted.

There is also an ambivalence (multi-valence might be a more accurate term) in the society with regard to what might be expected from education. Four, interdependent factors, which have important implications for expectations and in turn evaluation are: 1) student grouping, 2) the time period in which specified objectives are to be achieved, 3) kinds of objectives, and 4) who determines the objectives. Another phenomena which exacerbates the problem is that public education has evolved into a pattern in which students are aggregated into educationally manageable groups; given access to education services for varying lengths of time; required to pursue objectives designed to promote productivity, protect both student and society, and serve the twin functions of mental stimulation and liberation; and are held accountable for objectives which are mainly of societal origins.

Norms for gifted are nonexistent. Norms for achievement of gifted students do not exist for several reasons. Standardized measuring tools, by virtue of the very requirements they must meet to provide good information about students clustered near the mean, fail to provide adequate information for students at the extreme ends of the scale. The large numbers of students required for standardization cannot conveniently (or economically) be gathered from the sprinkle of gifted students to support adequately the norming process. Moreover, the diversity of exceptional abilities among the gifted create many areas of achievement thus further fragmenting any possible norming population. Teacher-made tests are more useful but suffer somewhat from similar problems. There are too many dissimilar gifts and too few students to make the benefits worth the effort.

Superficially, special provisions for gifted students seem much the same but there are differences which have important implications for evaluation. Grouping alternatives have developed which include, special attention in the regular class, clustering several students in a regular class, partial pull-out, a setting in which the regular and special curricula are integrated, and putting students together in congregated settings such as special classes or special schools. Time required to complete various activities is another obviously multi-valued parameter. Responding to needs of the gifted, as such needs become obvious, requires more than ordinary attention because progress is usually faster, there is greater need to respond to emotional dysfunction and their career planning matures earlier. Finally it has become increasingly more apparent that gifted students can effectively set objectives and criteria for themselves which are far superior to those set by well-meaning adults who do not understand the speeded developmental patterns.

One point of view is that it would be convenient and tidy if universal expectations existed in the education of the gifted, that is, if standards were established for each program component along with norms for student performance. If this were done, programming could more often focus on groups of gifted students. In the present state of the art this is not possible.

What of the future? Attempts to establish reference marks for comparison purposes require considerable time and effort. One or more generations of gifted students, having been given special educational programs to meet their needs, could, perhaps, provide the population and data for a longitudinal study which might help to establish standards for both program and student outcomes evaluation. This would help to establish expectations. Although some important benefits to students are observable in the short term, current knowledge does not yet yield a complete answer even to the short term value of special programs for gifted students, nor by what standards the programs themselves should be judged. The goal of making educational provisions which will enable the gifted to realize their potential for self-actualization and contribution to society is anchored in a future not readily available for scrutiny. In this circumstance it becomes important for evaluation to consider not only the achievement of objectives but also the worthwhileness of the objectives, and the provisions made for generating and revising them. In programs for the gifted, the evaluative processes shift sharply away from the usual, often statistical, comparisons to criteria which allow performance to be judged on a more divergent basis; a base that has its roots in the openendedness inherent in the multiplicity of students' gifts.

Some abbreviated consideration of philosophy, definition, identification, and curriculum are included in this chapter for such assistance as they may provide in establishing relevant expectations. A generalization can be made here. The most obvious effects of the differences between regular education and special programs for gifted students on evaluation, is a much expanded domain of evaluation and an increase in the use of less quantifiable data. Evaluation questions take less for granted. Data become more subjective (in the technical sense of that term).

## PHILOSOPHY

Statements of philosophy upon which the education of gifted students is based, serve the same purposes as in education generally, and very often are specific elaborations of the more general ones. The more useful statements appear to be those which provide reasons for the program, an indication of who is to be served, and some indication of the nature of the service to be provided. Philosophies usually originate from some mix of research, experience, and preferences for what ought to be.

This is true also in programs making provision for the education of gifted students however the proportionate contributions of research, experience, and preference make a very great difference. What is different is that research on educating the gifted is recent and meagre, experience is limited, and preferences are far too controlling. Yet it is this conventional wisdom, or presage data, which seems to determine curriculum decisions.

Presage data may be thought of as the sum total of applicable conventional wisdom. This term as used in this chapter and those which follow, denotes the accumulation of knowledge from both experience and experiment together with knowledge of the context in which it proved useful by those who make the decisions. No one person has all of it. Neither is it static; its application to either experience or experiment serves either to confirm it or to provide a basis for its modification. Using presage data is an art rather than a science; what data are applicable, how they should be weighted, and their applicability in the new context choice are matters of judgement or artistic choice.

Statements of philosophy. These statements often emanate from those who are experienced, observant of behavior, and have an depth of presage data. These statements can serve new programs, such as education for gifted students, in particularly important ways. They provide: a) a focus for program support, b) a direction for program development, c) a unifying force helpful in ensuring consistency in the program as it is developed, d) a backdrop against which the program may be judged, and e) a foundation contributory to program permanence. Programs which are based on, and consistent with, a widely accepted philosophy will most often have wide support in principle for the period of time necessary to demonstrate their worth.

Philosophy statements in education provisions for gifted students encompass new purposes and increase emphasis on established ones. An example of a new purpose is found in a school in Canada which intends to develop Olympic class athletes in some specified sports. An example of a new focus is implicit in such commonly used phrases as "self-actualization" being replaced by "realize their potential to self and society". While such a general philosophical statement may seem to differ little from that expressed (or taken for granted) in regular programs, purposes and efforts implied by them differ greatly. Enabling gifted students to reach their potential requires an increase in diversity of purposes pursued over a longer time frame. Enlargement of purpose demands elaboration of objectives statements and new efforts in curriculum to achieve them.

Consideration of objectives and curriculum concerns is deferred to subsequent sections of this and other chapters, but the expanded role of a philosophy statement in educating the gifted indicates the general appropriateness of the following kinds of evaluation questions which illustrate the enhanced scope of program evaluation in relevant programs. Program purposes and objectives become points of reference for evaluation but the requirements of a particular evaluation will naturally determine the actual questions developed and pursued.

Evaluation Questions About Philosophy. The following sample questions indicate the expansion of the evaluation question domain. As a statement of intent, is the philosophy

(a) commonly understood by:
  - parents?
  - teachers of gifted students?
  - other teachers?
  - administration?
  - board?

(b) supported by:
  - parents, and the community generally?

(c) useful as a guide to assessing consistency among other program components including:
  - definition?
  - identification?
  - purposes, goals, objectives?
  - curriculum?
  - evaluation processes and provisions?

(d) an adequate basis for obtaining support of the program in its developmental stages?

(e) comparable to other statements currently in use, having regard to:
  - the purpose?
  - indication of kind of service to be provided?
  - target population?
  - usefulness in encouraging support?

## DEFINITION

Most definitions in current use constitute an elaboration of the purposes to be served and a general statement indicative of who will receive the services. In essence the definitions are an extension of philosophy. The definition recommended by Alberta Education is a good example,

> "Gifted and talented pupils are those who by virtue of outstanding abilities are capable of exceptional performance. These are children who require differentiated provisions and/or programs beyond the regular school program to realize their contribution to self and society. Children capable of exceptional performance include those with demonstrated achievement and/or potential ability in one or several of the following areas:
> a. general intellectual ability
> b. specific academic aptitude
> c. creative or productive thinking, and
> d. visual and performing arts."

While the philosophical components should appropriately be given consideration, a definition addressing who is to be served fills another unique function; it is the bridge to identification. It therefore is a key element in assessing program consistency because one must address the appropriateness of curriculum in relation to clients served.

Evaluation Questions About the Definition. The following questions about definition could be included and expanded.

(a) Is the definition consistent with the intents stated in the philosophy?
(b) How does the definition compare with others in use?

## IDENTIFICATION

Identification of gifted children has received enough attention in current literature and in practice so that some general guidelines or principles emerge. Both research and experience reveal the fallacies in using a single instrument to identify gifted children. The result is agreement that no single instrument is suitable for identifying the variety of giftedness which may be expressed in many ways. Expression which reveals giftedness, is much more varied than the observed behavior of someone providing acceptable responses in the limited mode and scope of any test. For example, a very commonly used instrument is an individual intelligence test such as the WISC-R or the Stanford Binet.

Despite some extravagant claims to the contrary these tests do not measure such things as creativity, interests, motivation, or indeed more than a small number of the many components of intelligence.

While common usage and theory demands multiple instruments and observations, in practice most of the weight is often placed on one individual intelligence test. This practice is universally condemned in the literature, and some states in the U.S.A. have legislated against use of a single instrument for identifying gifted students, emphasizing by fiat the fact that none is valid by itself for the purpose. Furthermore, one-time use of a single measure is well known to be unreliable. Thus the first general principle is that identification must be based on multiple sources of information.

The use of too few data sources occurs usually in the interests of administrative efficiency. This suggests the second general principle. A school system should develop for its own use, a relatively small set of data sources found to be adequate for its purpose.

Various groups of people appear to have various rates of accuracy in identifying gifted children. It has been found that teachers who have not had specific in-service preparation for the task, miss about half of the gifted students in their classes. Parents are more accurate. Peers in grade levels six and up identify more accurately than do teachers without specific in-service preparation. This suggests the third principle. All teachers, inclusive of classroom teachers, taking part in selection procedures should seek appropriate in-service education.

The fourth principal also stems from the above information. Input into the identification process should include relevant information from parents, teachers, peers, and from the candidates themselves.

Data sources for identification. A number of data sources have been found useful in the identification process. These include results of intelligence tests (both individual and group) combined with tests of various aspects of creativity, biographical inventories, information about achievement, information about variety and kinds of interests, and about learning styles. Regardless of the instruments used, however, they are to be applied with due care because the usefulness of any measure depends upon its validity for the intended use and on the consistency with which it is administered. This suggests a fifth principle. Tests used in the identification process should be screened carefully to ensure validity for their intended use.

Principle number five has as its corollary a sixth. The identification process should be completed under the direct supervision of someone with a dual professional familiarity, with the use of test instruments, and with the nature of giftedness in children.

In addition to using various instruments of an objective nature, important information relevant to identification is available from parents, teachers, administrators and others. Information is neither acceptable just because it is gathered by objective instruments nor suspect because it is subjective in nature; the essential requirements are its accuracy and relevance. From this, principle seven is derived. The identification process appropriately includes both objective and subjective information relevant to the process.

The last principle is a result of the previous ones. The final selection of gifted children to receive special curricular provisions should be the responsibility of a group of people rather than one person.

## PRINCIPLES OF IDENTIFICATION

1. Identification must be based on multiple sources of information.

2. A school system should develop, for its own use, a relatively small set of data sources found to be adequate for its purposes.

3. All teachers, inclusive of classroom teachers, taking part in selection procedures, should seek appropriate in-service education.

4. Input into the identification process should include relevant information from parents, teachers, peers, and from the candidates themselves.

5. Tests used in the identification process should be screened carefully to ensure validity for their intended use.

6. The identification process should be completed under the direct supervision of someone with a dual professional familiarity with the use of test instruments and with the nature of giftedness.

7. The identification process appropriately includes both objective and subjective relevant information.

8. The final selection of gifted children to receive special curricular provisions should be the responsibility of a group of people rather than one person.

---

Relevance of identification principles. The relevance of these general principles for evaluation is that they constitute a tentative set of criteria with which the identification component of a program can be compared. Alternatively, if the principles are not acceptable as a standard in a particular evaluation exercise they may, nonetheless, be useful as a starting point from which the evaluator and the person(s) requesting the evaluation may develop appropriate expectations for comparison purposes. The principles as stated are not, of course, evaluation questions. Before they can be so used they must be reworded as questions and elaborated into appropriate subquestions.

Depending on the evaluation questions, evaluators may also need to consider appropriateness of personnel and/or organizational structure for implementing the identification criteria and procedures adopted. They should also consider the currency of knowledge held by teachers and consultants and the means for keeping abreast of developments in the field. The fidelity with which identification procedures implement the intent of the system's philosophy and definition of gifted children should also serve as sources of important evaluation questions.

Evaluation questions about identification. The following questions about identification are suggested for a beginning point.

(a) Are the above noted principles, or whatever principles have been developed by the school jurisdiction, being applied?
(b) Are those students identified by the procedures, members of the intended target population?

(c) Are any students in the intended population missed?
(d) Are those involved in identification adequately prepared?
(e) Are procedures administratively sound? Are they fair?

## CURRICULUM (Special Provisions)

Providing curriculum for gifted student programs is easily the most difficult of all of the program components. Developing a curriculum for this group of exceptional children may be described as an enigma. Teachers' comments indicate they recognize the pervasive problems and the work involved, but that they are also interested in the challenges. Curriculum building is difficult and often frustrating because the need for new challenges exists but teachers feel they lack the knowledge and understanding necessary to do the job adequately.

The provisions which a teacher (or school, or system) makes to meet the challenges of "curriculum fit" do not have any close counterpart in regular programs. They therefore constitute an extension to the domain of evaluation questions in programs for gifted students. Evaluators may rightly consider for assessment not only the special curriculum presented, but provisions for developing, deciding, and implementing it. Making such provision is complicated by the unusual difficulties in developing special curricula for the gifted. There are four complicating factors: a) the nature of giftedness itself and the giftedness of children so identified, b) the setting for curriculum development, c) deficits in teacher preparation, and d) articulation with the regular curricula. Brief descriptions of these factors and their implications for evaluation will indicate some enlargement of traditional evaluation concerns.

---

### COMPLICATING FACTORS OF SPECIAL PROVISIONS

---

1. The nature of giftedness itself and the giftedness of children so identified.
2. The setting for curriculum development.
3. Deficits in teacher preparation.
4. Articulation with the regular curriculum.

---

## The Nature of Giftedness Itself and the Giftedness of Children so Identified

In essence, gifted students are different in their exceptional abilities from their more average peers to such an extent that their needs are not accommodated in the regular curriculum. Gifted students learn faster, think in greater depth, explore more widely, discover new questions to which they seek answers, and according to some indications many think differently. But on the other hand they may also suffer some negative effects such as having problems coming to terms with their own giftedness, having low self-esteem and self-confidence and unsatisfactory relationships with peers and adults. Nonetheless, even prior to school entrance, the energies of many have been selectively applied to the world about them with the result that their understanding in areas of particular interest usually exceeds by several years that of average students, while at the same time many have also become uncomfortable with their differences from age peers. Gifted students also differ markedly within their own group. They vary in interests, patterns of exceptional ability (gifts), self-understanding, a sense of self-worth, self-confidence, maturity, ability to channel their energies, willingness to accept the fact that they are different from age peers, the opportunities they have had for their own development, in short, in just about every conceivable way. They are more different from each other than from the average and although they all appear to have quicker mental processes than their age peers, they differ also from each other in this regard.

There are implications for evaluation stemming from these differences from age peers. It is of extra importance that those involved in evaluating programs for gifted students be aware of the unique characteristics of gifted students which affect the nature of the education provided to them.

> FAMILIARITY WITH THE NATURE OF GIFTED CHILDREN IS NEEDED BY THOSE CONDUCTING AN EVALUATION IN ORDER TO SUCCESSFULLY CARRY OUT THE PROCESSES OF PROGRAM EVALUATION AND STUDENT EVALUATION.

The level of awareness should, as a minimum, include recognition that: a) gifted students have special "needs", and that some of the more important ones are in the affective domain, b) the needs of gifted students are not assessed in routine school procedures and are only partially revealed in the identification process, thus requiring additional attention, c) meeting the needs will require some degree of invention in curriculum content, delivery processes, and settings, and flexibility in selection and use of personnel, d) these unusual requirements on teachers call for more than providing ordinary assistance.

## The Setting for Curriculum Development.

Programs for gifted students are relatively recent additions having their setting in schools with well established curricula taught by teachers whose accepted role is to interpret and implement that curriculum in rather uniform physical surroundings and with a plethora of tested resources. The regular curricula which they implement have been developed and approved by other people in other arenas. Traditionally they have been developed by departments of education using a variety of forums and with the assistance of experts including curriculum specialists, psychologists, and subject matter specialists. These curricula have been developed over time periods measured in years if not decades and with expenditures of hundreds of thousands of dollars. Curricula for gifted students, on the other hand, are most often decided, developed and implemented by their teachers who have neither a public forum, extended experience, nor long periods of time in which to try out, and find acceptability for, the curricula they create.

Presage data is widely used in planning and operating education programs for gifted students. It is the basis of all of the curriculum content and activities provided, of identification procedures, of the organization for curriculum delivery. Because of its nature, application of presage data requires: a) care in securing an adequate base of presage information. Since no one has all of the relevant presage data, securing an adequate base may require input from several persons or groups and from other sources. b) an element of creativity. The application of presage data is invoked precisely because the treatments-outcomes correlations are inadequate; there is not sufficient information to identify the treatments. Creation of a set of treatments is in order. Assembling and reconfiguring concepts, processes, etc., is the universal creative process. c) after the fact assessment to determine whether the treatment was successful or not. If so, the presage data is confirmed. If not, some modifications in the current wisdom, i.e. presage data, may be suggested by the experience.

A major implication for evaluators is the need to understand that developing the "special provisions" is nearly always a task for teachers of the gifted. It is usually not a traditional curriculum, though it may have its beginnings in traditional content and traditional settings. The activities are developed to meet the needs of children with special talents and abilities who are often disregarded. The curricula are often developed by children with leadership from the teacher and are so individualized that what might appear as one curriculum to the casual observer are really as many curricula as there are children in the classroom. The effort in developing these curricula and their applicability are often overlooked in the evaluation process.

## Deficits in Teacher Preparation

Rarely do undergraduate teacher preparation programs include courses on education of gifted children. Teachers, moreover, have not generally been prepared by either professional education or experience for curriculum building. Their preparation continues to be directed toward implementing curricula rather than developing it. At best the process which teachers of the gifted adopt involves gathering a library of resource materials, selecting, adapting, and organizing them, then putting them in usable form for delivery. Sometimes, as might be expected, teachers simply adopt someone else's curriculum. What is most often lacking in either process is the preliminary rationale defining purposes and objectives, and the subsequent cycles of tryout, assessment, and revision to ensure that the curricula achieve the objectives and/or that the objectives themselves are worthwhile. While widely known and capable educators have applied their energies to building curricula for gifted students with some useful results, the extreme abilities of the student and situational variability make prepared curricula less appropriate than teachers of gifted students might like or the intended use demand.

Deficiencies in teacher preparation are not confined to curriculum construction. Understanding the nature of giftedness and gifted individuals, developing objectives, creating suitable curricula, and devising and conducting delivery strategies and activities, constitute only part of the challenge. A truly remarkable flood of information, opinion, and speculation about developmental stages, brain hemispheric functioning, learning styles and preferences, and teaching styles is stirring education to its depths. All of these movements have more importance for teachers of gifted children than for other teachers who normally are much less deeply involved in curriculum development or with gifted students for whom the new information is more crucial.

Deficits in teacher education have implications for evaluation. The potential domain of evaluation is, once again, extended to questions about availability of resources and suitability of processes available for curriculum development. To the extent that curricula do not meet student needs then the reasons must be sought. The teacher is one resource and if he/she does not understand the nature of giftedness nor how to adapt curricula to meet the needs of the gifted in the classroom, the opportunity for student learning and challenge is a risk. Since programs are often evaluated on the basis of student achievement and student satisfaction, the teacher must be seen as one of the key figures.

Teacher preparation is a new program component deserving consideration in planning an evaluation. A major U.S. study found that teacher in-service is the greatest unmet need in educating the gifted, thus identifying another important extension of the evaluation domain. Evaluators should be asking questions about the quality of learning opportunities for both students and teachers.

## Articulation With the Regular Curriculum

The peaceful co-existence of special provisions for gifted children along with the regularly prescribed curriculum presents special problems. Although the fundamental principle on which provision of special education for gifted students rests is that the regular curriculum is not appropriate, students are still expected to satisfy its requirements. To what extent the regular curriculum in regular classrooms can be easily adapted to meet the needs of the gifted has been the focus of considerable study and some debate But what else needs to be done, and whether the regular classroom setting is adequate also requires attention. Integration within their own classrooms, partial segregation, and complete segregation all have proponents and advantages as the setting for curriculum delivery. Integration is said to maintain a sense of reality and contact with the world in which the gifted will live. Working with groups of gifted peers helps the gifted put their own exceptional abilities in perspective, thus generating self-knowledge and confidence. Each setting also has its own problems: extra demands on teacher time, scheduling, and developing a composite curriculum. What part of the regular education experience may be foregone so that a gifted student or group of students will have time and energy for the special experiences deemed appropriate is another complex problem.

Planning curriculum for gifted students is, even more than is usual in regular curriculum development, an exercise in futuristics as it tries to identify and address student needs. Yet the regular curriculum also tries to address current and future needs. The worth of special provisions for the gifted will be decided ultimately by observation of the effects upon lifetime careers not yet decided upon. It is much too early to determine if it can it be done effectively by teachers in their current setting. A host of curriculum material, concepts, ideas about the special needs of gifted children, and special objectives are being developed in such things as creative problem solving, critical thinking skills, higher level thinking processes, lateral thinking skills, divergent (productive) thinking skills, independent learning skills, philosophy for children, understanding their own special gifts, risk taking etc. But a smorgasbord of interesting activities may bring satiation without satisfying students' future needs for life and career skills.

Much of the debate about which type of curriculum will do the most good for students is not easily answered. It leaves open the opportunity to do some longitudinal research comparing the various curricula and the results over an extended period of time. It also leaves open the opportunity to compare the long term effects of curricula on those who are allowed entry to the special provisions for the gifted and those, though close to the cutting score, were denied entrance.

Providing curricula for gifted students is an important, difficult, and never ending task of very considerable importance. Program evaluation therefore must pay particular attention to provisions for developing, implementing and evaluating curricula. For the evaluator there are a number of caveats. 1. The curricula should be examined in the light of what is possible under the circumstances in conjunction with current wisdom about what is best for such students. 2. The provisions and structures for constant re-examination of curriculum offerings are vital and deserve to be given thorough examination. 3. The way in which teachers of the gifted operate should be examined to see whether there is the flexibility and openness required in the curriculum provided to the students. In addition to discovering what works with gifted students the teachers should adopt an investigative, evaluative, and explorative stance in looking for what might work better and what the outcomes ought to be in both the pragmatic and philosophic sense. There are few guidelines to help develop a scope, sequence or content for gifted students oriented as it must be toward careers and personal interest which are years in the future. 4. Resources to accommodate structure and energy to encourage exploration, reconsideration, and revision of curriculum may be just as important as the ongoing services extended to teachers in developing their own. 5. Curriculum delivery will necessarily diverge from regular classroom practice in order to implement properly the differentiated curricula.

Evaluation Questions About Curriculum (Special Provisions). The questions that could be asked by evaluators concerning special provisions are many, varied, and dependent upon individual circumstance. These circumstances include needs of children, numbers and physical setting. The following questions are only a suggested beginning.

(a) How do the special provisions differ from the regular curriculum?
(b) What special skills and knowledge are developed or addressed?
(c) How is the regular curriculum addressed?
    - what parts are left out?
(d) How are students grouped for instruction?
(e) Are the special provisions sufficiently flexible and open to meet student needs?
(f) In what way are the special provisions meeting needs of the students?

## SUMMARY STATEMENT

While this chapter is about the state of expectations in programs for the gifted, the expectations suggested are tentative as befits the state of program development and the understanding of student achievement. The reader should understand that suggestions are tentative and should be carefully selected and used only to that degree which has validity for the intended applications.

# CHAPTER THREE

## PROGRAM EVALUATION

Evaluation is commonly talked about as a necessary aspect of the educational process to ensure that goals and objectives are being addressed. Yet no single model in common use seems to provide adequately for evaluation of programs for gifted students. The literature too, points to this fact, as does the evaluation project of Clarke and Nyberg (1985). It was interesting to find that attempts to use some of the models in various evaluation activities within each of the three school systems sponsoring this project, led to the same conclusion.

Unfortunately many systems lack program evaluation service capabilities and some of those having such services evidence difficulties when evaluating programs of a special nature such as provisions for the gifted. Some evaluation proposals appear to consist of mixtures of various models and approaches, or to employ data gathering instruments which are not appropriate to the specific program components being evaluated. Some methods used seem to be oblivious to the questions being asked or the kinds of data required to answer the questions.

What is even a more regrettable situation is the fact that all too often programs for gifted children are developed without any thought as to determining whether they meet expectations. Some programs hang on year after year just because they are put in place, whether useful or not. Other endeavors disappear, as if by whim, even though some components appear to be worthwhile and meet special needs. The diversity of programs for the gifted is easily recognizable and each one has unique problems for an evaluation team. But while the instruments and protocols developed in one system may not easily transfer to another, the effort to evaluate should not cease.

There is the usual nervousness about evaluation of programs for the gifted that is endemic to all new programs. There seems to be a fear that evaluation will bring a decision to remove something that required great effort to initiate. For some reason the concept of evaluating to determine where improvement can be made is foreign in many programs for gifted children. The mere existence of the program appears to be proof of its worth. Yet there is disquietude in the educational community and among the clients of educational pursuits, that programs are not adequately meeting needs and that improvements are slow to surface.

## EVALUATION CONCERNS

Alberta Education has put school jurisdictions on alert by requiring evaluation at all levels from board to student. This has raised some concern about the directions evaluation of programs for the gifted might take: what will be in jeopardy because of summative scrutiny? and what will be the shape of education as a consequence of the formative evaluation procedure? There is concern that these provisions may be swept away prematurely before the formative evaluations have time to make them viable or that they may become confused or compromised with summative evaluation purposes. Another concern is that new program components such as identification and selection which are not part of regular education, will be neglected. The possibility that evaluation may overlook providing adequate support to teachers, such as in-service education and the services of support staff, is also of concern because the requirements of teaching the gifted are unusual.

### Formative Evaluation or Summative Evaluation

When the main purpose of evaluation is to provide a basis for improvement of a program it is called formative evaluation. It is the kind most commonly invoked in education, particularly for assessment of new and innovative programs in which even brief experience almost always reveals previously unsuspected faults and suggests improvements. As the more obvious faults are removed and improvements implemented, assessment procedures become more formalized and sophisticated, and the program is revised and developed into a more finished and final form. But whether simple or sophisticated any assessment which has program revision as its purpose and/or effect is a formative evaluation.

Summative evaluation is defined by one specific purpose and/or effect, namely, a decision to continue or discontinue the entire program. It is possible to apply a summative evaluation procedure to part of a program. In such a case the distinction between summative and formative becomes blurred, but is resolvable on the basis of the size of the part of the program under consideration and the strength of opinion about its significance to the entire program. When the part of the program being considered for elimination is the major fraction of program effort, or when it is an essential program element, summative evaluation procedures are appropriate because in either case a decision to discontinue would have the effect of terminating the entire program.

Although most evaluations in educational practice exhibit both purposes, i.e. to improve on or to make a decision about continuance, to some extent, the procedures appropriate to the two kinds of evaluation will likely differ in a number of significant ways. These differences are important, even critical in the evaluation of provisions for gifted students, even though the two procedures are superficially quite similar. Some important differences follow.

1. The locale of decision making will likely differ. In summative evaluation a political decision is probably appropriate; in formative evaluation professional staff will be involved, if not solely, then at least more substantively.

2. Evaluation questions will differ. In summative evaluation the focus is on product. The basic question is whether or not the product, i.e. the totality of outcomes from the program, justifies the effort. This will include all outcomes whether intended, unintended or, as may well be the case in programs for gifted students, identifiable as desirable outcomes only after the fact though not prespecified or even prespecifiable. In formative evaluation, the spotlight is on process and the cause-effect relationships between processes and products. Summative evaluation must consider the totality of products. Formative evaluation may, and most often does, address the processes and their effects in a more piecemeal fashion. The process of identifying gifted students, for example, may be the main, or the only, area of concern in a formative evaluation. Within that area, the purpose may be even more limited to comparing the effects of using one instrument with an alternative one.

3. Data forming the bases of decision making will differ. In summative evaluation the actual and/or perceived total worth of program outcomes (products) is paramount. In some fashion this worth must be weighed against the cost of

the program. It is quite likely that opinions of persons or groups with vested interests will be among the most persuasive data components, especially as these opinions reflect political influence. Expert opinion will also, of course, be useful but often to a lesser degree than in formative evaluation operations. In formative evaluation, information which provides detail about the processes in the program and which relates process to outcomes is of central importance. Data about alternative processes and related outcomes become the major bases for the evaluative comparisons which shape the programs. Data about costs of various alternatives is also an important base for decision making. In formative evaluation the main thrust is to make the processes, and hence the program, more effective and cost efficient.

4. Reports differ. Summative evaluation will usually require a more complete report, since it is more likely to be in a setting with adversarial overtones where process, findings, and conclusions are all more likely to be under attack than in formative evaluation. The reports differ also in another very significant way; the audiences are different in kind. As noted above, the important audience for a summative report is probably political in its intentions, while for the formative report it will most likely be one with professional intent.

5. Frequency of application is different. In current practice in evaluating programs for gifted students, formative evaluation will constitute 95% or more of the total evaluation effort.

The literature on formative/summative evaluation contains argument in favor of combining the two forms, suggesting that a formative evaluation could be the basis, completely or in part, for a summative one. While there is merit in that contention, the advantages are heavily outweighed by the confusions and redundancies which would result from attempting to serve both purposes at once; namely, to select questions which endeavor to screen data sources on two criteria, to answer differing questions for two quite different audiences, and to prepare distinctively different reports. In assessing the provisions for gifted students where these differences have greater significance, advantage lies with separation.

For all of the above reasons and considering also the newness of the programs it is recommended that formative and summative evaluation procedures be kept separate and that the major evaluation effort continue to emphasize the formative needs of programs for gifted students.

SOME DIFFERENCES IN FORMATIVE AND SUMMATIVE EVALUATION

| Item | Formative | Summative |
|---|---|---|
| The locale of decision making. | Political decision. | Professional decision. |
| Evaluation questions. | Focus is on process. | Focus is on product. |
| Data forming the basis of decisions. | Detail of process and relationships of process to product is central and determines alternative actions. | Worth of products, programs, outcomes important. Broader base of information is necessary in addressing outcomes. |
| Reports. | Detailed report for a professional audience. Formative report is necessary for the summative report. | Detailed report is necessary for a political audience. |
| Frequency of application is different. | Currently takes 95% of the effort in programs for gifted children | Currently takes 5% or less of the total effort in evaluating programs for the gifted. |

## EVALUATION STRATEGY

Rather than attempting a piecing together from various models, the approach chosen in doing the evaluations in the three school systems in the project was more fundamental and generic, starting with the evaluation-question concept and proceeding through various stages to the final one of reporting. The experience resulted in a strategy which appears to solve a majority of the problems and allays some of the concerns; it is recommended as a valid approach.

The program evaluation strategy which emerges from the mix of literature gleanings, experience, and tryout is a set of seemingly simple procedures which, despite their apparent simplicity, require considerable care in application. The strategy in its elemental form calls for the following sequence of procedures.

- develop evaluation questions
- identify appropriate data sources
- develop appropriate data gathering procedures

- organize data
- answer evaluation questions
- report answers

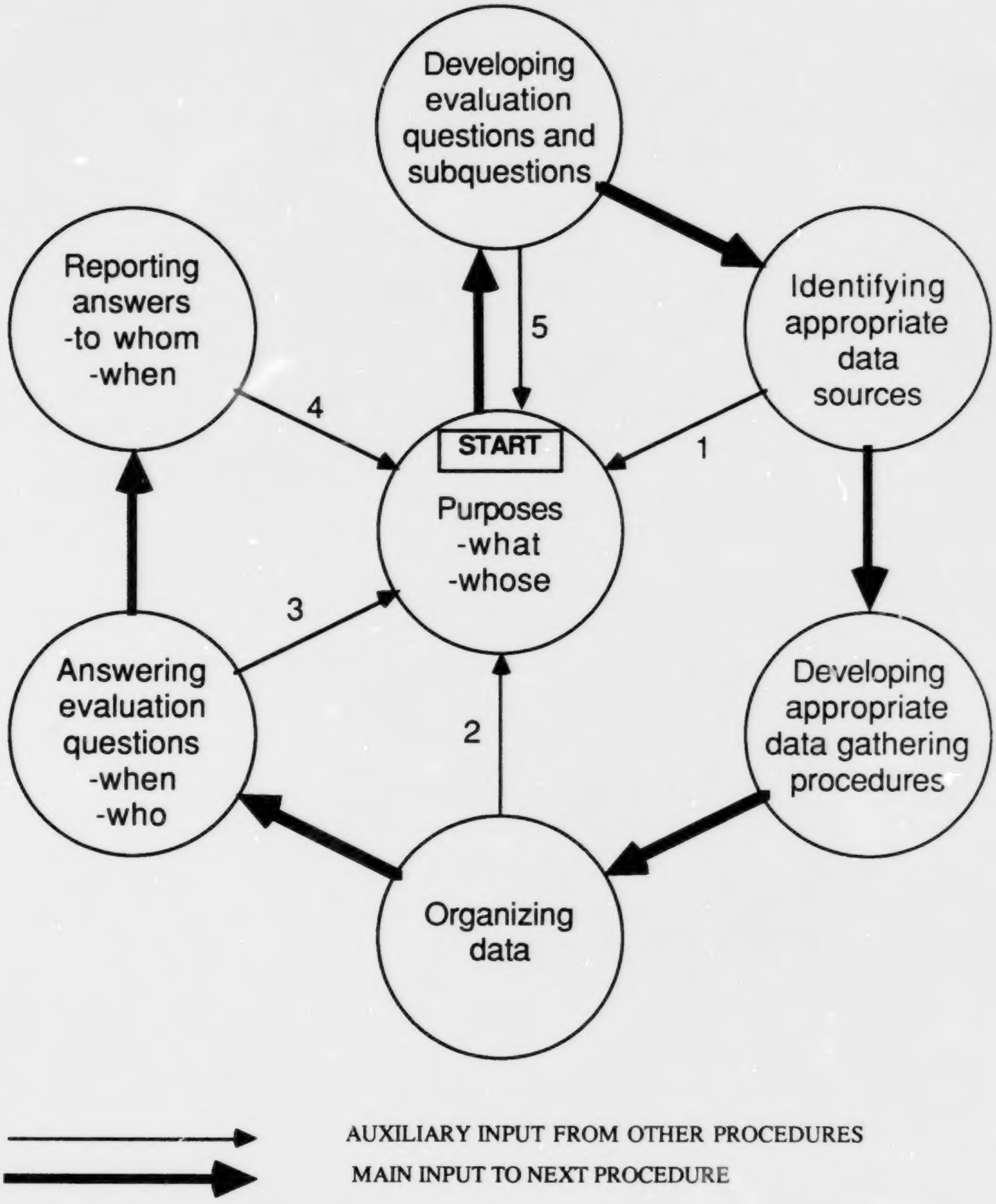
Each of these six procedures is more complex than it might appear at first sight and each is an interlocking part of the whole evaluation. The following subsections explore this complexity and relatedness.

Implementing the Strategy. The strategy is essentially a configuration of procedures in a time sequence, with the outcome of one procedure providing the input for the next. The heavy arrows indicate the logic of the output-input sequences. The starting point is a statement of purposes (which, itself, is an output of some pre-evaluation process). The purposes statement becomes the input for developing evaluation questions, and the developed questions, then, become input for identifying appropriate data sources, etc.

Although the main input to any procedure is made from the immediately preceding one, other important inputs come from the remainder of the process (and possibly from other sources). The most common of the components which affect the final statement of purpose are indicated in the diagram (Figure III-a, page 27) by the light arrows. These are numbered and the input represented by each is identified as follows:

1. Data available may modify purposes, suggest extension or limitation.
2. While organizing data affects reporting directly, purposes must also be kept in mind.
3. Pilot phase: answering question may reshape purposes.
4. Purpose statements may become more explicit by timing and audience and the pilot phase.
5. Purposes must be adequately represented by the evaluation questions.

Figure III - a


Developing evaluation questions and subquestions
Reporting answers -to whom -when
Identifying appropriate data sources
START
Purposes -what -whose
5
4
1
3
2
Answering evaluation questions -when -who
Developing appropriate data gathering procedures
Organizing data
AUXILIARY INPUT FROM OTHER PROCEDURES
MAIN INPUT TO NEXT PROCEDURE

1. Availability of appropriate data confirms purposes. If data are not available, modification of evaluation questions or statement of purposes may be necessary.
2. Organization of data may reveal possibility of extending or modifying purposes.
3. Check to determine if purposes must be modified because one or more questions are not answerable.
4. Timing and audience information will provide specificity to purpose statements.
5. This forms part of the feed forward-feed back loop to check that the purposes are adequately represented by the evaluation questions.

## SEQUENCE OF PROGRAM EVALUATION PROCEDURES

I. Develop Evaluation Questions
II. Identify Appropriate Data Sources
III. Develop Appropriate Data Gathering Procedures
IV. Organize Data
V. Answer Evaluation Questions
VI. Report Answers

---

### I. Developing Evaluation Questions.

This process is the base of the evaluation and is paramount in importance. Each of the remaining five procedures, as will become apparent, depends materially on this one. This interrelatedness places strict requirements on the questions. There is the necessity for questions to be lucid, cover the concerns of those whose needs are to be served by the evaluation, and be limited to those concerns, while serving as a basis for gathering sufficient and appropriate information. These criteria help to ensure focus and economy of effort.

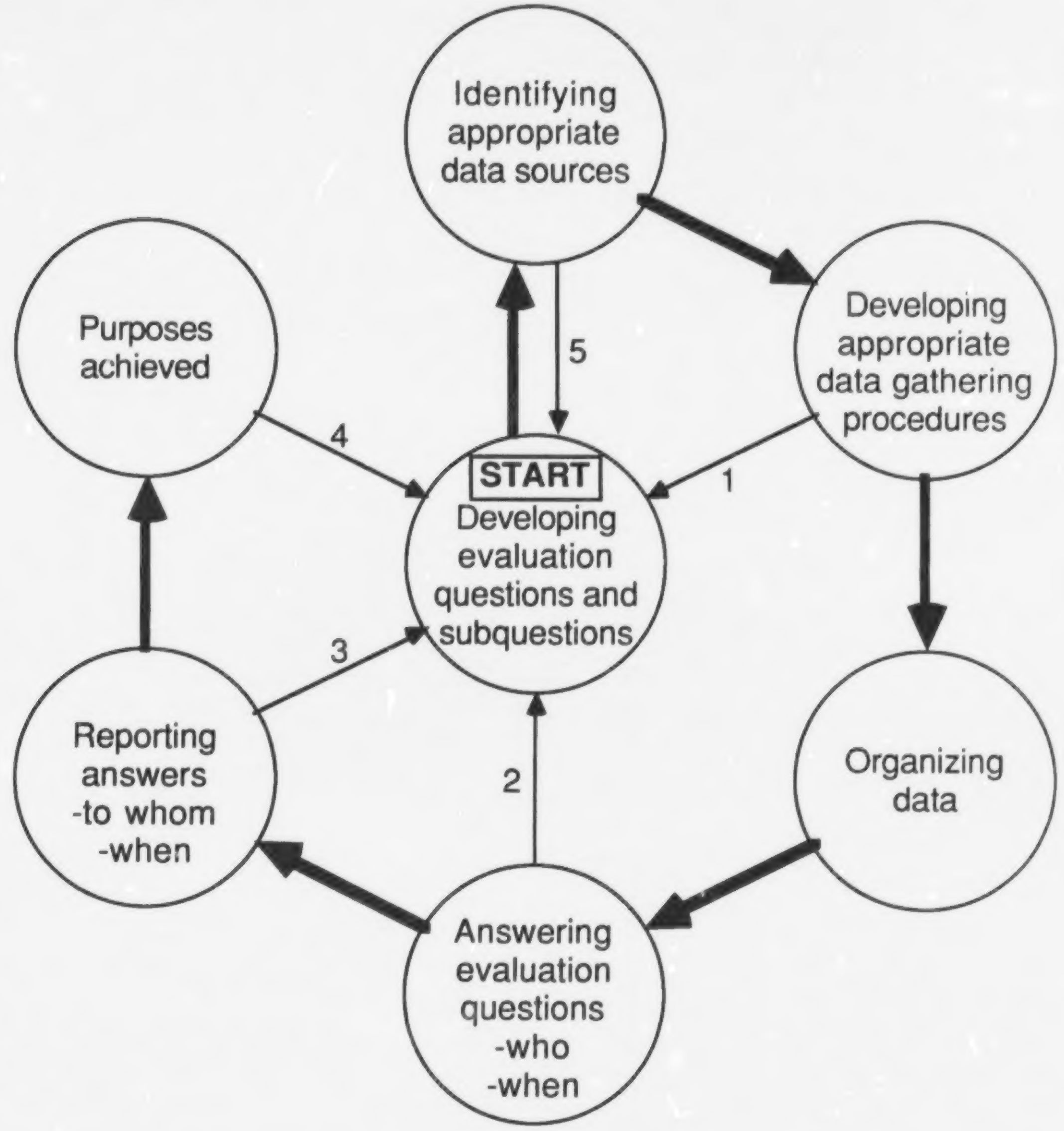
Each of the other parts of the process provide valuable information to that piece at the focus of attention. (Note Figure III-b, page 29). The information gleaned during consideration of the data gathering operation could modify the questions being developed. Information from the pilot phase, as well as the timing and reporting requirements, will also impact question development. The set of purposes to be served by the answers is relevant also and may assist in determining the sources of information.

A similar diagram could be used for the four remaining procedures of the evaluation strategy by placing each one, in turn, at the center of the diagram and noting the influence of the remaining parts.

Developing evaluation questions includes a number of subprocesses.

1. The first step is to accumulate and articulate questions of interest, particularly those which are important to persons or groups having a vested interest in the program and with an unsatisfied curiosity about its merits. Qualifying as having such an interest, and very

Figure III - b


Identifying appropriate data sources
Developing appropriate data gathering procedures
Organizing data
Answering evaluation questions -who -when
Reporting answers -to whom -when
Purposes achieved
START
Developing evaluation questions and subquestions
1
2
3
4
5

AUXILIARY INPUT FROM OTHER PROCEDURES
MAIN INPUT TO NEXT PROCEDURE

1. Information from data gathering procedures, such as instrumental/procedural validity, and effort required, may suggest revision to questions.
2. Pilot phase efforts to answer evaluation questions may indicate necessary changes in the questions.
3. Timing and other reporting requirements may require modification of questions.
4. Positive feedback here (in pilot phase) indicates viability of questions as a basis for evaluation.
5. Part of the basic feedback loop which will insure that data exist for answering the evaluation questions.

often a legitimate one, are those involved in developing or operating the program for gifted students: teachers, administrators, boards of education, parents of gifted students, and sometimes the students themselves. This list is obviously not complete nor ordered in any particular way. It is usual, however, for the person in charge of developing the program to be the most interested and curious, possibly because the experience is still a pioneering one and therefore the position is a lonely one, lacking in feedback. In this initial process of developing evaluation questions, various degrees of formality ranging from a request for input from interested parties to group sessions of these parties, may be invoked. Broadening the base from which the key evaluation questions arise will usually improve the probability that real concerns will be addressed and that answers to the evaluation questions will serve a useful purpose.

2. Select those questions which are to be specifically addressed by the evaluation. Since more valid and interesting questions can be raised than it is feasible to answer, priorities must be established. This may be done by: a) setting priorities to the purposes the evaluation is to serve, b) assigning each question to the purpose(s), and c) assigning a weight to each question according to its importance in achieving the purpose(s) to which it is related. The final selection of questions should be deferred until data sources and gathering procedures are under consideration, since availability of data is often a decisive factor in selecting evaluation questions.

3. Elaborate the evaluation questions into sub-questions. Nearly every question of any substance needs to be broken into sub-questions (and often some related questions) in order to identify possible data sources and facilitate data gathering. This process is both analytic and creative, and so is not amenable to a simple set of directions. The product is, however, recognizable a) by its usefulness in identifying the data needed for addressing each of the sub-questions, b) by satisfying those whose purposes are most directly served, that the overall question will be answered by answering the sub-questions, and c) by satisfying the evaluator that the sub-questions are answerable and that they provide an adequate basis for answering the evaluation question. An example, taken from the evaluation process used in Strathcona County is provided as an illustration.

## DEVELOPING EVALUATION QUESTIONS

1. Accumulate and articulate questions of interest to all who have a vested interest in the evaluation.
2. Select those questions which will be specifically addressed.
3. Elaborate the evaluation questions into sub-questions.

Evaluation Question: Is the Challenge Program effective in meeting the needs of gifted students?

Elaboration:

a. Are there procedures in place for assessing student needs?
    i. Through identification procedures?
    ii. Through other procedures?
b. Is the information used to determine the program?
c. In what ways does the Challenge Program meet the needs of students in the program?
d. In what ways does the Challenge Program benefit other students?

It is obvious that item "d" in the "Elaboration" is an extension of the evaluation question. It was added because some students identified as gifted were not in the challenge program and some other provisions were to be made for them. It was also determined in the initial interviews with program consultants that other students had benefitted in a variety of ways. Items "a", "b", and "c" constitute an elaboration of the evaluation question which was acceptably answered on the basis of the response to the three sub-questions. Clearly, other elaborations would have been possible, and more detailed elaboration of "c" might have been needed. In the actual evaluation as carried out, 'acceptability' of the answer was predetermined on the basis that if student needs were being identified, and if the needs so identified were used to decide program for the students, and if confirmed benefits to students related to their needs, the question could be answered satisfactorily by the evaluators in consultation with the program director. Judgement by these people, of course, demanded much more than a simple 'yes' or 'no' to the questions, as posed. Such questions are, however, the major guides in the logic of answering the questions and serve as priming to the analysis of what data is to be gathered and how it is to be used.

## II. Identifying Appropriate Data Sources.

Appropriate data have three essential characteristics. They must be: a) accessible (available), b) accurate (reliable), and c) useful for the purpose to be served (valid). In the program evaluation process they must provide an acceptable basis for answering evaluation question(s), i.e. valid for this purpose. They are also characterized by what purpose and whose purpose it is.

Availability. The accessibility of data is best determined with the assistance of someone quite familiar with the school system and the special provisions made for educating gifted students. Some data may be immediately available for gathering, such as a teacher's record of curriculum activities, or an anecdotal record of student achievements, or student logs, while other data may require development. It is not always easy to ascertain information about creative problem solving skills, ability to function in the higher levels of Bloom's taxonomy of cognitive skills, or quality of classroom interaction. For such data, availability depends on the time and effort required for its development. Availability is a first screening for appropriateness. Information about what data are available and the effort required to collect and develop them is of course preliminary to the actual processes of selection and collection and should be part of the pilot process.

Accuracy (reliability). The reliability of the data is essentially the responsibility of the evaluator with assistance from the person(s) who has a fundamental working knowledge of the system. Reliability is of concern in all procedures relating to identification of data sources and data gathering. Some sources are inherently more reliable than others, i.e. are more accurate representations of fact or more accurate measures of a "true value". Written records tend to be more reliable than recollections. But recollections may serve to enhance records in providing depth and in some cases go well beyond what was recorded, thus providing new information. There is always the danger that recollections can be imprecise.

In evaluating programs for gifted students, recollections may be the only source of some important data. Bias often interacts with facts in memory and thus less involved participants tend to be more objective and reliable. More often, however, what was found to be available, were recollections about the same events by several persons having different perspectives. Tapping several sources tends to improve accuracy by confirming some data, revealing inaccuracy in other data, offsetting biases one against another and at the same time improving the depth of understanding obtained. As might have been anticipated a combination of written record and recollection was found to be a more accurate and useful, in-depth representation of data than either one by itself. This process has been called triangulation.

Another general kind of data is inherently reliable. When people express their feeling or give their opinions, the data provided is reliable, unless by intent the opinion given is inaccurate or is inadvertently miscommunicated. Evaluation of programs for gifted students has relied somewhat heavily, perhaps too much so, on opinion. Collected opinion, however, if carefully done, is a useful, reliable, and valid indication of such things as satisfaction, confidence in what is being done, or sentiments about what should be done.

Some commercially available instruments may be useful in providing reliable information to the user. In using these instruments, the reliability of data obtained will depend on careful administration of the instruments and is limited to the specific kind of application for which the instrument was constructed. The evaluator, however, is very likely to be faced with the task of building many more instruments than can be found suitable for her/his use, and their reliability becomes an immediate concern. Measures of internal consistency such as the split half, or KR20 tests for reliability, may be used if the extensiveness and nature of the information collected thereby are such as to make their use appropriate. Where such measures are appropriate, they provide a generally recognized indication of reliability. In the evaluations associated with the project, however, there was very little opportunity for their use, because they were not appropriate or simpler checks on reliability existed.

The interview, one data gathering process used extensively in this project, was found to have several advantages in obtaining reliable data. In a field marked by imprecision in usage and understanding of some commonly used terms, the interview allowed the researcher to penetrate the confusion and pinpoint the exact information needed. It also gave a more sure indication of when a data source was depleted. A further advantage noted in one system was the uncovering of some important but latent evaluation questions.

The following four devices were used to improve reliability of both data and data gathering processes.

a. Assuring anonymity to the extent desired by those supplying information improved the accuracy of the data collected.

b. Making specific the statements of data required, its sources, and how it was to be accessed. (Specificity is probably the evaluator's best friend). Precise and explicit statement of the data to be gathered appeared to lead to improved accuracy of the data gathering instruments constructed and the procedures used. For example, in gathering data to answer a question such as, "Are the procedures, which have been approved by the system, used in identifying gifted students?", the researchers were able to break down the question into more precise specifications including system intentions, instruments to be used, when they were to have been applied, by whom, and the expected decision making processes. This analytic approach was helpful in collecting accurate (and valid) data. It also helped decide when the data source was exhausted, relieving in some measure the pressure on respondents "to fill in all of the blanks". A more general statement of the above question, "Are identification procedures appropriate?" could also be answered by similarly precise statements of the data to be gathered, when, and by whom, providing such specification could be made. In such a situation, however, the specific statements would need to be developed and then submitted for acceptability, by those whose purposes were being served, in order to ascertain their usefulness or validity as a basis for answering the evaluation question.

This process is essentially a way of attempting to set up comparison criteria in a particular situation. It will often fail. When this happens, recourse to expected outcomes, or to current wisdom (presage data) about identification procedures must be taken.

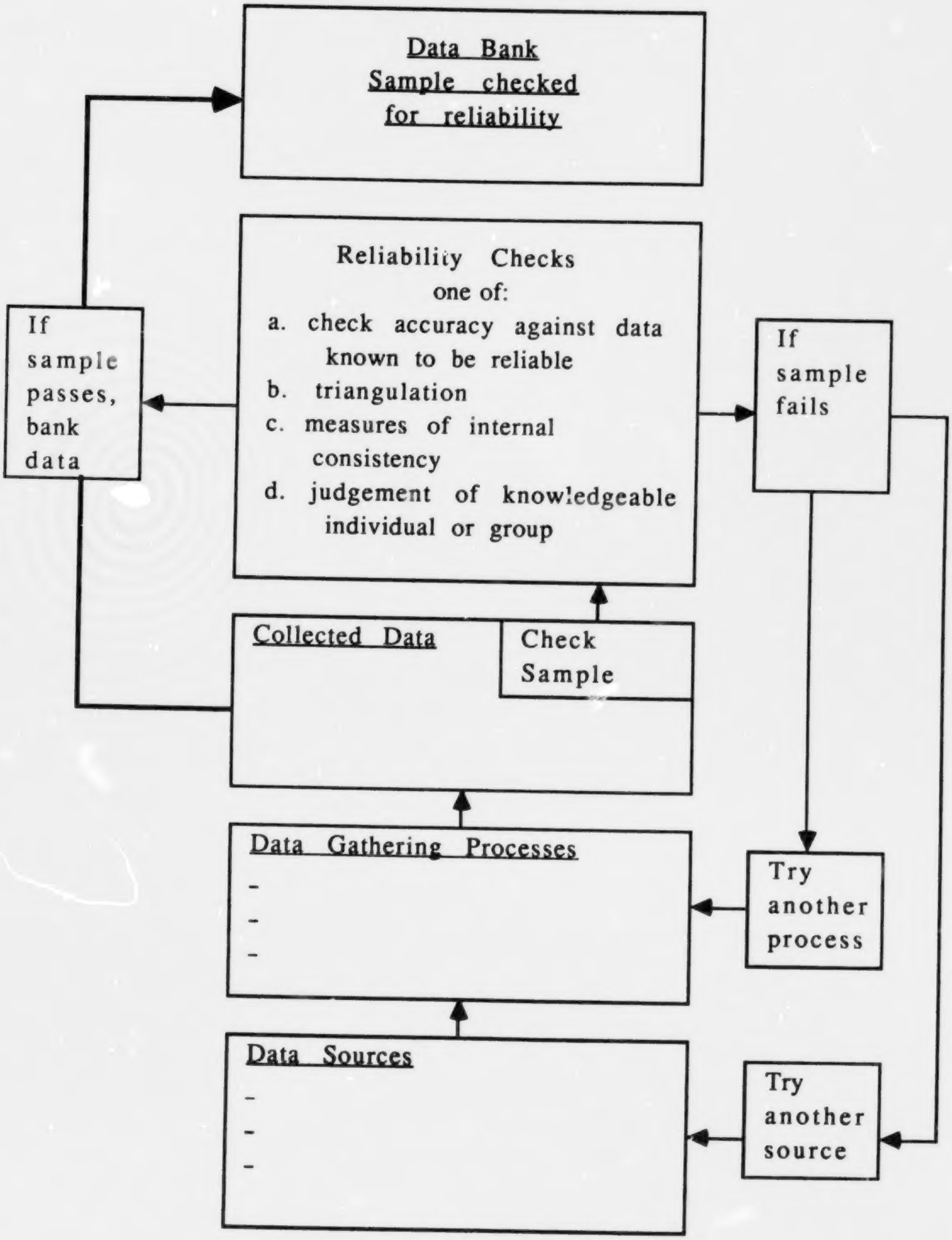
c. Giving careful and adequate attention to piloting data identification and collection. In the program evaluations conducted as part of the project, more than the usual attention to piloting the data, gathering and developing protocols, and testing the reliability of the data obtained, was found to be useful and necessary. There is a tendency to shorten the piloting phase as it may seem to be wasting time, but adequate attention to piloting will assure accuracy and thoroughness.

d. Broadening the base for judging reliability. Triangulation was used as one way of doing this. Another was use of judgement panels of two or three knowledgeable people to assist the evaluator in considering data and deciding whether it was adequately accurate.

These steps for improving accuracy of data are necessitated in part by circumstances. When instruments are developed for use in evaluating programs for gifted students, it is usually on a one time in one system application and so the opportunity of deciding reliability through repeated measures or repeated applications of the same measures is usually not available, although, as already noted, measures of internal consistency can sometimes be used as a check on reliability.

The diagram (Fig. III-c) illustrates application of the reliability quality control measure suggested above. It is a simple feedback loop from data sample, through the reliability check, to the data gathering process, with an appropriate branch to the reliable data bank or back, as needed to revise either the process or source. Repeated cycles without achieving reliable data status suggest that the data source (or process) may not produce reliable information and that an alternative source/process should be explored. Like availability, reliability is a necessary (though not sufficient) condition for data usefulness (validity).

Figure III - c
Reliability Check


Data Bank
Sample checked for reliability
Reliability Checks
one of:
a. check accuracy against data known to be reliable
b. triangulation
c. measures of internal consistency
d. judgement of knowledgeable individual or group
If sample passes, bank data
If sample fails
Collected Data
Check Sample
Data Gathering Processes
-
-
-
Try another process
Data Sources
-
-
-
Try another source

Notes: Reliability, although mainly a pilot phase concern, should be checked occasionally throughout the data gathering process. Sample should be representative of data but need not be random. It is chosen on the basis of whether one or more of the reliability checks can be made. Reliability is a necessary but not a sufficient condition for validity.

Validity. This is crucial and is generally defined in terms of its usefulness for the purpose(s) to be served. Speaking strictly, no process or instrument has validity unless its use contributes satisfactorily to the purpose to be served. One consequence is that validity can only be determined by the user who has a specific purpose in view. A second consequence is that validity must be determined on location, i.e. in a particular school or system with specific purposes in view. Hence validation is an integral part of the evaluation process.

The evaluation strategy suggested in this chapter, requires that the data base must be useful (satisfactory) for answering the sub-questions. It also means that each sub-question must satisfactorily serve its intended purpose in answering the main questions which in turn must serve the intended evaluation purpose(s) of an identified person or group. If the purpose is formative, for example the answers must assist decision making designed to confirm or improve programs. When on the rare occasion the purpose is summative, the answer must illuminate that decision.

Whatever the purpose, the process of validation is always a necessary component of evaluating education programs. This is particularly so for new programs with as many unusual features as those for gifted students. In these programs, the processes necessary to ensure valid data should never be ignored or treated superficially. Fortunately the often difficult task of securing acceptably valid data is facilitated by the evaluation strategy suggested.

Attention to the linkages from data to purpose as noted above, is the basis of quality control for the validation process. In the following representation of the process, each box represents an evaluation process outcome and each arrow represents a set of one or more questions to check whether one outcome provides an adequate basis for satisfactory achievement of the next outcome. The outcomes and questions, along with suggestions for who should answer the validity check questions, are shown in Fig. III-d (page 38).

Although many, and perhaps most of the validity concerns in evaluating programs for gifted students are attended to by the quality control checks indicated in the chart (Fig. III-b), some are not. In essence the process indicated is a check on the usefulness of data gathered and on the instruments used in the data collection process. The intent and the expectation from the process is assurance that the data will be valid for the specific evaluation purposes.

## Figure III- d
## Quality Control: Validity

| Evaluation Process Outcomes | YES | Validity Check Questions (all of the following necessary) | NO | Check Made: by whom? when? |
|---|---|---|---|---|
| Data | | For answering in each subquestion, are data: | | |
| | | a. avaliable? | revise subquestion | Ev., X, SE P. 1-4; E. 4,6 |
| | | b. suitable? | find new source | Ev., X P. 2-4; E. 4,6 |
| | | c. sufficient? | expand data bank | Ev., X, SE P. 2; E. 2,4,6 |
| Subquestion answers | | For answering each main question are subquestion answers: | | |
| | | a. a suitable basis? | re-check validity | Ev., X P. 1,2,3; E.2,6 |
| | | b. a sufficient basis? | expand subquestion set | Ev. P. 1; E. 1 |
| Main question answers | | Are main question answers a suitable basis for: | | |
| | | a. making intended decisions? e.g. about curriculum, identification, organization, etc. | review and revise main questions | Ev., X P. 1,2,6; E. 6,7 |
| | | b. completing other aspects of intended purposes? e.g. reporting, for new program development | re-examine/ reorganize data | Ev., X, SE P. 1,2,6; E |
| Evaluation purposes completed | | | | |
| Reports completed | | | | |

Code:
1. Developing evaluation questions
2. Identifying appropriate data sources
3. Developing appropriate data gathering procedures
4. Gathering data
5. Organizing data
6. Answering
7. Reporting answers

P. Pilot phase
E. Evaluation phase
Ev. Evaluator
X. Person or group whose purposes are being served
SE. System expert

The validity thus obtained is called content validity. (Availability and reliability of the data foundation are both necessary elements of validity and may be checked concurrently with the checks on content validity indicated in Fig. III-d).

Two other kinds of validity are more troublesome. The first is construct validity, the second is predictive validity.

a. Construct validity. This is of concern where the questions or sub-questions deal with constructs, i.e., abstractions such as creativity, giftedness, self confidence, sense of self worth, intelligence, and leadership. The basic problem with constructs for our purpose is that, like impressionistic literature or art, they connote more information than they convey and the connotations vary from reader to reader. The main application of construct validation is to resolve this problem using sophisticated analytic procedures, such as factor analysis, to sharpen the meaning of the constructs and increase the commonality of the connotations. Evaluation of programs for the gifted and evaluation of student development, are not suitable settings for this kind of validation, each of which requires sophisticated statistical procedures. But evaluation questions and assignments often use the terms for the abstractions, thus creating problems and confusion for unwary evaluators.

The remedy is to avoid using abstractions by specifying a set of observables as replacement for the abstraction. For example, everyone concerned might agree to replace, for the purposes of the evaluation, creativity with the set: {fluency, elaboration, flexibility}. The latter are observable, even measurable and can represent the term creativity in the evaluation. The problem in using such a set of observables to represent the construct is not its use in program evaluation, but in the almost irresistible tendency for everyone, the evaluator included, to overlook the awkward set of observables and use the more convenient but less accurate abstract term. "Creativity", for example is not equivalent to the above set nor any set yet created. It usually means different observables (and some things not yet observed) to different people and consequently confusion results. The persisting confusion is pandemic to the field of constructs. Energetic attention for over two thirds of a century, for example, has not completely clarified the construct called intelligence. However, since use of sophisticated research procedures to clarify the constructs will probably never be one of the procedures demanded in evaluating programs for gifted students, construct validity, though of interest, will not be a serious practical problem.

b. Predictive validity. This is much more challenging. Its simple purpose is to relate treatments (curriculum activities, program organizational arrangements. . . . ) to effects (student development, program efficiency, . . . ). Its measure is in the strength and reliability of the correlations between treatments and outcomes. Consideration of its application to evaluation in programs for gifted students, however, is more effective as a vantage point in identifying problems than as a source of solutions to them.

If correlations between treatments and achievement of the major goals of a program for gifted students could readily be ascertained, a very important set of evaluation questions would be answered, questions about effectiveness of the educational provisions made for the student.

Finding the correlations is not easily done for a number of reasons. First, the major goal of education for the gifted, [to enable students with] "outstanding abilities . . . capable of high performance . . . to realize their [potential] contributions to self and society", is usually recognized only by its achievement after an extended period of years during which the "treatment", or enabling educational experiences must have been provided. During this time, so many experiences are provided that it is usually impossible to identify those which cause the results. Second, individual potential is itself, an elusive construct; the potential of the multi-capable gifted student is even more elusive. Furthermore, the "potential" is not constant over time but is known to wax or wane on the basis of school and other experiences of the student contributing further to its elusiveness. Thus the "worth" of the abilities and skills developed in the individual by the experience cannot be measured against his/her "potential". Third, there are no norms for gifted student achievement, hence no

> comparison standards. Fourth, individuality among gifted students is so much the rule it is hard to amass sufficient data to obtain reliable correlations between treatments and outcomes.
>
> It eases the problem of establishing correlations somewhat if the long-term goals are broken down into short, medium, and long-term objectives. It becomes easier to relate treatments to effects in the shorter time periods. But on the debit side is the need to relate short, medium, and long-term objectives to the long-term goals; this is more than a slight complication.

All kinds of valid data are appropriate for program evaluation. Educators generally prefer to rely on objective information and shun the subjective. But in evaluating the benefits of special experiences provided to gifted students a reversal is often necessary. Because of the 'ceiling effect' observed when standardized tests are used with gifted students, and also because these students are usually offered a greatly extended variety of learning experiences, the appropriateness of these measures becomes much less certain. Such instruments are always to be used with a degree of caution, but, particularly in the special setting of evaluating outcomes for gifted students, validity may be dubious and sometimes nonexistent. For some purposes no such measures are available. Teacher-made objective tests may be so constructed as to avoid the "ceiling effect" and are likely to be valid representations of student achievement, provided appropriate performance criteria can be determined. This is a difficult task. Increased reliance must, consequently, be placed on subjective ones.

---

## ESSENTIAL CHARACTERISTICS OF APPROPRIATE DATA SOURCES

---

1. Availability
    a. someone familiar with the system.

2. Accuracy (aided by)
    a. assuring anonymity
    b. making specific the statements of data required, its sources, and how it is to be accessed.
    c. piloting data identification and collection.
    d. broadening the base for judging reliability through triangulation and judgement panels.

3. Validity
    a. construct (deals with abstractions not having readily defined boundaries).
    b. predictive (relates treatments to effects).

---

In an attempt to reduce subjectivity, a variety of checklists and rating scales are often used. Data, whose accuracy may be inconclusive from one source, are often verified by data on the same point from one or more other sources; a procedure called triangulation. Similar considerations apply to other program components as well, because standard expectations and procedures fit imperfectly and sometimes do not fit at all. These considerations suggest wide exploration of data sources prior to deciding which to tap.

Data sources pertinent to an evaluation may include:

- educational literature, both general and specific, including research and expert opinion relating to the education of gifted students.

- documents held by various agents such as teachers, schools, system administrators and consultants, school board members, provincial education department officials.

- information and opinions held by students, parents, school based administrators, system administrators and consultants.

- student logs (not uncommon in education provisions for gifted students).

A final step before deciding which sources are appropriate for each sub-question is to examine a sample of data available from each of the plausible sources. This step has the obvious purpose of preventing unpleasant surprises which could arise if the source, though apparently appropriate, yields inadequate data. Making the judgements as to which source to use, is the responsibility of the evaluator and is best done with adequate awareness of the program, the purposes for the evaluation, the persons requiring answers, potential data sources, and with due consideration of the kinds of data actually available.

Reference to the elaboration of the evaluation question used in the preceding section may be useful. First the kind of data required was decided and then the source was identified using the following format.

As may be observed from the example, the kinds of data range from documented information to information which is only accessible directly from the various categories of participants.

## III. Developing Appropriate Data Gathering Procedures.

Responsibility of evaluator. The task of developing appropriate data gathering procedures is essentially the responsibility of the evaluator. Paucity of usable ready-made instruments, lack of established procedures in evaluating programs for the gifted, and the extensive variety of activities within almost any given program, make the task challenging. The usual caveat, minimal interruption to the program, must of course, be observed.

Scope of evaluation must fit questions and data sources. The task of developing appropriate data gathering procedures is unusually difficult because new instruments and processes must be generated, which take into account the increased variety in program elements such as objectives, student identification, assessment of student needs, and curriculum alterations. Building new data gathering instruments for a specific, application (usually one time) in a specific school or school system is vastly more simple than that facing experts in building instruments for multiple purposes in many settings. A test prepared by a teacher in an evening, for example, may well be a more valid instrument for her purpose than any commercially available test, although the latter may have taken experts several years to build. So also the data gathering instrumentations required in a particular program evaluation application is a reasonably manageable task.

Gathering Data. The procedures must be appropriate to the setting and will likely vary with the extensiveness of the special provisions. It may be appropriate to develop questionnaires which will be sent to every participant including district and school administrators, consultants, regular classroom teachers, teachers of the gifted, students, and parents of both students in the program and those who do not have children involved. It would also be appropriate to select a sample of the participants to interview as a check on the validity of the information and also to determine if more relevant information can be gleaned. In small jurisdictions, the entire population of participants might receive questionnaires, whereas in larger school jurisdictions only a representative sample will be used.

Another data gathering procedure often employed is the visit to the classroom. These visits need to extend across several days rather than be only a single visit of a few minutes or even one day. The visit should include an opportunity to discuss the intents of the teacher and student and observe whether these are being met by the special provisions activities. The usual qualifications of an evaluator in a classroom setting must apply, i.e. they must be knowledgeable about the nature of giftedness and have familiarity with appropriate curricula and methodology. These visits could be supplemented with data from tests, student products and other appropriate information such as questionnaires or checklists which measure the teaching and learning environment. It is necessary to be cautious about observing only the teaching and not the teacher.

In some cases only an interview system might be used. In this case the evaluation questions must be developed in advance and quite strictly adhered to, though there should be an opportunity for those interviewed to give unstructured comment. Once more, the size of the school jurisdiction and the evaluation resources available will determine sample size.

As part of the data gathering procedure, evaluators should address the collection of relevant documents from the school jurisdiction, individual schools, and information in literature which address the purposes of the evaluation. These documents and information which express philosophy and intents of the program are of vital importance and provide, in large measure, the basis for determining program success. It might be found, however, that there are other standards, such as student and parent pleasure, which become the determinants of success.

Collecting relevant information is a prerequisite to answering the evaluation questions and can take many routes as noted earlier. In the fieldwork which formed one of the bases for this book, participant interviews supplemented with questionnaires were the main means of accumulating the appropriate information. Examination of system documents, (school and teacher) use of simple survey forms, and classroom observation accounted for most of the remainder. Other methods of gathering data such as the use of checklists, opinionnaires, standardized and informal tests, etc. prove useful if accuracy and relevance can be maintained.

Monitoring Data. The use of interviews is a good means of monitoring data, provided it is done by someone familiar with the information required and the purpose to be served. An alert interviewer can assess relevance and accuracy of the information until satisfied that these requirements are met, or that no usable information is available. Specifying with precision the information to be collected proves to be more important than exactness in how questions are worded because flexibility in approach is often needed. Gathering information by means of interviews is usually more certain, and although more time consuming for the interviewer, may be less so for the person interviewed, than collecting the same information by other means.

The prompting questions from the interview protocols which were used with challenge teachers, students, and parents in gathering information to answer item C) of the elaboration example, illustrate some of these points.

C. In what ways does the Challenge Program meet the needs of students in the program?

Parent interview item: What are the benefits to your child from the Challenge Program?

Student interview item: What are the benefits to you? what activities give these benefits, are there disadvantages ?

Teacher interview item: What are the benefits of the challenge program to the class as a whole? to individual students? to which student(s)? and what benefit(s)?

Triangulation. An illustration of the triangulation process, spoken of earlier, is apparent in this example; it consists of comparing the responses of parents, students and teachers. Although it is not always possible to key the response sets exactly, enough such comparisons were made in this instance to validate the benefits claimed.

Questionnaires sent or delivered to the sample of participants must be precisely worded and will usually be received more favorably if concise. This means that much effort must be expended in developing the appropriate questions for each of the participants. Even though the general question to be answered is the same for more than one group, each question must be worded to fit the perspective and experience of a specific group so that there is clarity of understanding and no room for misunderstanding. This is the major reason for piloting questionnaires and making appropriate revisions prior to sending them.

When other data gathering procedures are used, as much or even more care should be taken to ensure accuracy and relevance of the information obtained. Data collected by one process is often validated through another, such as interviewing. Data gathering is complete when the evaluator and the person whose purposes are being served thereby, agree that the information is adequate to answer the questions.

---

## DATA GATHERING PROCEDURES

---

1. Responsibility of evaluator but should have minimum disruption to classroom.
2. Scope of evaluation must be adjusted to fit the evaluation questions and data sources.
3. Data sources should be selected so as to validate information through triangulation.
4. Data gathering methods:
    a. questionnaires, surveys, ad hoc instruments
    b. interviews
    c. classroom visits
    d. district and school documents
5. Monitoring data for relevance and accuracy.

---

## IV. Organizing data.

The basic requirement here is that the information needed to answer each sub-question be assembled in a convenient way. The procedures of identifying data sources, developing data gathering procedures, and gathering data, normally have the effect of dispersing the information, particularly if, as is usually the case, each source provides information relating to several sub-questions, and/or several sources contribute to a given sub-question. In the example given above, the information from parents, students and teachers about benefits to students, would be related to item 4 c. The necessity for this kind of organization is discussed in the following section.

## V. Answering evaluation questions.

An evaluator of programs for gifted children may find that answering the questions based on the information gathered may be simple, complex or even sometimes next to impossible. For some sub-questions the collected data make the answer obvious, requiring nothing further except to state the answer. An example of a simple question for which the data makes the answer obvious is, "Are system approved procedures for identifying gifted students used in the schools?" For other questions varying degrees of judgement based on amount and variety of experience is needed; for still others the evaluator will need the services of other experts. A question requiring expert opinion might be, "Are the system approved procedures adequate?" Finally, for some questions, there is no readily available answer because there is a lack of research evidence and experience to address the issue.

It is generally recognized that the evaluator should be experienced in the field being evaluated, but some specialized aspects require greater expertise. The evaluator's task is to match the expertise necessary with the question being answered. Since several people may be involved in answering the questions, there is a need for assembling and organizing the data to make it comprehensive but also confined to the questions under study.

## VI. Reporting answers.

Reporting evaluations of programs for gifted students is much the same as for other education programs and similar conditions apply. The basic considerations are timing, recipients of the report, content of the report, and method of communication.

Timing. The timing requirements are usually quite obvious. The report is required before decisions requiring the information are made. These decisions may be either formative or summative. It happens, not infrequently, that an incomplete report with tentative answers must be provided rather than a completed one; the latter only being available after decisions have had to be made. It is suggested in the literature, and was found advisable in carrying out the evaluations, to give some preliminary indications of findings. This means the report should be circulated to the supervisors or overseers of the project in draft form to receive their input and reaction. Two purposes may be served: first it provides opportunity for additional information to be advanced which might qualify the findings, and second it provides adjustment time for those who might otherwise react inappropriately out of sheer surprise and chagrin.

Recipients. Who receives the various answers from the report depends on the purposes to be served. What the various recipients require should of course be provided, plus those additional answers which are otherwise useful or interesting if there is no specific reason not to do so. Very often all of those who have a legitimate interest can receive the entire report making it unnecessary to prepare multiple segments.

Content. The content of the report will depend on the purposes to be served. This in turn determines the amount of data and detail to be included as supporting material for the answers.

Method of Communication. The form of communication used in the report can vary from completely verbal to completely written or formal. Verbal communication, where appropriate, is much quicker and often saves a great deal of time preparing written reports. A written report has the advantage of being able to be thoroughly checked over time, reviewed as occasion warrants, and even being consulted to ensure that recommended modifications are implemented. A mix of the written and verbal forms provides some saving in time as well as a record of the answers to the questions, hence evaluation reports are frequently a mix of both oral and written.

## THE WORTH OF A PROGRAM

Program worth is judged mainly by reference to proven benefits for students. Measuring such benefits is the subject of Chapter Four. This type of measurement is not a precise science. Identification; curriculum content, sequence, and delivery; program evaluation; and student evaluation all depend heavily on judgment involving measurement and presage data. In this situation, the processes involved in decision making assume great importance and, consequently, assessment of program decision making processes become an important part of program evaluation. Some of the more important program processes are: providing teacher in-service education, nominating and selecting gifted students, setting program goals, selecting and/or developing curriculum content, deciding curriculum delivery methods, using non school personnel such as mentors, parents and other experts. Processes which are self improving have obvious benefits over other routines. In the present state of the art in educating gifted students, it is a safe premise that developmental processes augur better for viable programs than do the products judged as "good" by current standards.

As an example of this focus on process, E. Susanne Richert has designed a checklist of the important elements in evaluating or setting up a program for gifted students. Her checklist is made up almost entirely of either explicit processes or alternatives lists which imply processes. The checklist is published in the Gifted Children Newsletter Vol. 55, No. 5, (May 1984), p. 18. It was taken from The National Report on Identification, (by the same author) available from the Educational Information and Resource Center, Box 209, Sewell, N.J. 08080.

## SUMMARY OF STRATEGY

The strategy discussed above is summarized to provide a concise and perhaps more convenient reference. Each of the six major steps is numbered with appropriate concerns listed in descending order. (A sample worksheet taken from the evaluation of special provisions activities at Calgary Board of Education is included in Appendix D.)

1. Developing Evaluation Questions
    (a) purpose of evaluation
        - must be intelligible to all
        - cover concerns
    (b) identify those who need to know
        - when
        - why
    (c) develop a pool of questions
        - accumulate and articulate possible questions
        - select appropriate questions
        - elaborate questions with sub-questions
    (d) review to satisfy evaluator and system personnel
        - questions must relate to concerns
        - sub-questions must be answerable
        - insure that answering sub-questions will answer main evaluation questions

2. Identifying Appropriate Data Sources
   - (a) must be available
   - (b) must be accurate
   - (c) must be valid; acceptable for answering questions according to purpose being served as well as who the purpose serves
   - (d) piloting question
   - (e) use of objective as well as subjective information
   - (f) judging reliability through triangulation and judgement panels

3. Developing Appropriate Data Gathering Procedures
   - (a) responsibility of evaluator
   - (b) scope adjusted to questions and data sources
   - (c) selection of data sources
     - use of one source to validate another
   - (d) gathering data
     - interviews
     - interviewer must be familiar with information
     - information required more important than wording of questions, allowing flexibility in approach
     - use of ad hoc instruments
     - examining documents
     - survey forms
     - observation
     - consistent monitoring of accuracy and relevancy

4. Organizing Data
   - (a) assemble in convenient form
   - (b) relate information to each particular sub-question

5. Answering Evaluation Questions
   - (a) answers sometimes simply need stating
   - (b) match expertise necessary with the question being evaluated

6. Reporting Answers
   - (a) timing
   - (b) recipients of the report
   - (c) content
   - (d) method of communication

# CHAPTER FOUR

## EVALUATING STUDENT ACHIEVEMENT IN SPECIAL PROVISIONS ACTIVITIES

Education programs may be evaluated in terms of efficiency, acceptability, appropriateness of content, perceived benefits to society, and numerous other features, but ultimately, the worth of any program of instruction must be decided on the basis of what it does for students. This is the key question in any educational evaluation including the special provisions made for gifted students. If it is not specifically part of an evaluation, the benefits are either seen as obvious or an assumption is being made that significant benefits are derived.

The authors of this report found, in visiting programs for gifted students, reported in a previous study (Sillito and Wilde, 1983), that evaluation of students involved in programs for the gifted seemed to be lacking in many instances. Teachers and administrators questioned on this point admitted that it was a neglected area. While they thought it was probably important to evaluate these students, they were unsure as to whether the gifted should be compared to students of similar chronological age and grade or compared to gifted peers. There has been indecision as to whether these students should be evaluated at all, and if so, on what basis, since they are at the top of the academic achievement scale anyway. It was even suggested by some teachers at one school involved in the current study, that to evaluate gifted students may act as a deterrent to them wanting to be involved, and this might ultimately result in destruction of the program.

During the course of this study, several discussions were held with teachers and children about evaluation procedures and the worth of evaluation. It was the consensus of opinion that evaluation was necessary, but the extent and appropriateness of the methodology were areas where opinions were inconsistent.

The attempt in this chapter is to outline the concerns about evaluation of gifted children as voiced and alluded to by teachers, provide reasons why evaluation of gifted students is necessary, supply some suggestions made by experienced teachers about how to evaluate these students, and give a brief statement about reporting results. Unfortunately, there are still far more questions and challenges than answers and solutions.

## PURPOSE OF EVALUATION

One reason for evaluation is to provide feedback to the learner on his/her progress so that performance or achievement can be efficiently improved. This is often overlooked by teachers and school administrators as they think about evaluation of gifted students. They fail to realize that a necessary part of the learning process is feedback on the adequacy of performance or achievement in relation to some set of predetermined criteria or even criteria developed while the project is in progress. While it is possible to use a single performance upon which to establish future criteria for judging adequacy, this usually occurs only with those who are at the very top of their performance category and where the expectations are previously established by someone else on some theoretical basis. This may occur with creative products and performances or with someone who establishes a new relationship among concepts and is therefore at the cutting edge of knowledge.

Unfortunately though many teachers do not comprehend the significance of evaluation in the learning process and it is probably true that most students are naive with respect to the purpose of evaluation as well. Students often see grades and comments merely as a teacher's assessment of performance, however derived. The end result, too often, is to view evaluation as summative rather than formative with respect to learning.

One could justifiably ask, "Why be concerned about evaluating gifted students?". Since these students are already at the top of achievement, why appraise their efforts?

Being gifted results in students having the capability of perceiving at a more mature level while still seemingly very young. This ability to perceive relates to the notion of evaluation just as much as to other concepts and constructs. Teachers can help the gifted child achieve insight into evaluation at a much younger chronological age than we normally expect of similarly aged peers. In addition these children seem to more readily understand that evaluation is necessary to improve learning as well as to establish worthwhile and appropriate evaluation criteria. These children want to be evaluated. While we should be striving to help all children understand the importance of evaluation in the learning process, it is easier with the gifted because of their enhanced perceptual capacity. This merely says to educators that what we can do with gifted children, we should also be doing with other students at the appropriate time in their educational career.

Gifted students, usually being more insightful, desire feedback about performance so as to make improvements. While this is not always so, it should be the responsibility of the teacher to be explicit on this point and demonstrate its importance to students by demanding that improvement be made in future performances commensurate with previous appraisals.

When teachers were asked what they thought was the purpose of evaluation, they agreed on the following:

1. to encourage further growth in a child,
2. to demonstrate achievement of a curriculum,
3. to develop the program and facilitate aims of the child (formative),
4. to advance teacher purposes, i.e. to show they have achieved something; a means of justification, even to "teacher peers",
5. to provide feedback to parents and students, and give a diagnosis of skill development, and
6. to facilitate research.

Even though not all of these apply directly to the child, it does show that teachers are aware of the need for evaluation as an aid to student achievement.

## CHALLENGES OF STUDENT EVALUATION

The problems of evaluation have an interesting genesis and a logical development. When the educational enterprise undertook to remedy the ills in the system for those with undeniably exceptional abilities, labelled gifted, an interesting diagnosis was made that curricula were to blame.

The result was the implementation of many innovative curriculum ideas. These changes prompted reconsideration of objectives, redefining of objectives and attempts to specify criteria upon which to measure success. In turn attention became focused upon the inadequacies of evaluation practices.

Traditional approaches to evaluation usually rely upon standardized achievement tests and teacher-made tests. The former assume normal distribution of traits in the population and develop norms based on a comparison of relative performance. The latter is also based on a comparison of performance except that the sample is much smaller and the test items less rigorously controlled. While standardized tests can be worthwhile in some instances, they tend to be an injustice to the gifted who are academically inclined, because these students cluster at the top without a sufficient spread of scores to make adequate judgements about superiority. This ceiling effect in standardized tests is also a result of the inadequacy of test items to measure the performance of the gifted. Teacher-made tests may suffer from the same problems and normally only teachers having sufficient experience with gifted children are able to make the most accurate judgements; this only occurring when they have developed criteria for that purpose.

The consequence is that not only are normal classroom expectations unsuitable as standards for the gifted, there are no norms for the gifted as a group. Two Alberta teachers involved in teaching gifted children summarize the problems and suggest direction in the following statement.

> Divergence in programs and differences in gifted children make the old testing tools ineffective. New data need to be acquired and codified; [they] must be relevant, must cover a wide spectrum of objectives, must have a manageable storage and retrieval system, and must be useful in assessing student progress.

In the course of the present study, there were many other important points raised by teachers. Objectives of the regular curriculum are also objectives of the gifted. Teachers noted the need to have measures of quality such as is needed to measure the interaction which occurs during discussion sessions. They went on further to say that evaluation is very difficult because each child is unique and there is no way to make adequate comparisons. Some teachers of enrichment classes, often at a loss as to how evaluation might be accomplished equitably, decided to give few if any evaluative statements. Unfortunately this procedure also seemed unfair and inadequate in the minds of gifted children.

When teachers were asked to express some of their concerns about evaluating gifted children, they made the following statements.

## TEACHERS' CONCERNS IN EVALUATING GIFTED CHILDREN

1. Standardized tests show some relationship to the rest of the district but they do little to help in the classroom.

2. Because curriculum must come from the children and not be imposed, we must work with their interests and even help the child discover interests when none are seemingly evident. This has many implications for evaluation.

3. These children are just as anxious about tests as any children. Too much pressure, whether from teachers, parents, peers or self, can result in breakdown.

4. Reporting evaluation is tiring. Often the evaluation is very subjective but the children want measurement against some standard and they want a thorough explanation of how their grade was determined. Students want to know why - they want the evaluation explained in detail and even want to be a part of the evaluation; this is where they differ markedly from the normal child.

6. Students want to know the standards ahead of time and these are not always available.

7. There are so many differences in the gifted - in language expression, content learning, thinking skills, etc.

8. Parents have some problems understanding the differences between classes for gifted children and those for the regular classes. It is difficult to satisfy them when they want grades to make comparisons.

9. How do you evaluate the most brilliant innovations of a child and reflect it in a grade score while taking into account that the child is not brilliant in all areas?

10. It is difficult to capture the total growth from a long time period and reflect it in a single grade.

11. When students are allowed to pursue 'strange' goals, it is difficult to evaluate them because it may be a block to their progress and remove the necessary challenge to continue the effort.

---

Summary Comments. Some general statements about the findings of the researchers seem to be appropriate as a summation to this section. Many teachers are left to their own devices when it comes to evaluation. There is a lack of consistency in the approach to evaluation within a district and within the school. Teachers complain that they do not know how to evaluate gifted students properly and are given little guidance even when they specifically request help. While there is more help provided in the larger systems, there is still much left undone and teachers are unsure as to the basis for evaluation and therefore struggle with how to provide feedback to the students on improvement. It was found that where teachers present Bloom's Taxonomy in an effort to teach children the thinking skills required, they also use this as the criteria upon which to evaluate the adequacy of any product. While this system has proven to be effective in some instances, many teachers expressed dissatisfaction because it failed to meet their needs.

## PRINCIPLES OF EVALUATION

Based on interviews and discussions with teachers of gifted children and also based on the literature, there seemed to be a number of conclusions about evaluation which could be expressed as principles. Most of these have grown out of experience with differentiating curriculum developed to meet the special needs of the gifted.

## PRINCIPLES OF EVALUATION FOR THE GIFTED

1. To the degree that students are given Individualized Education Plans (IEP) or Individual Program Plans (IPP), evaluation must also be individualized.

2. To the extent that objectives are not, or cannot, be defined by objectively decidable criteria, subjective measures must be used.

3. Where the curriculum is designed to serve a complex of objectives rather than atomic ones, the evaluation must similarly address the complex as an organic whole.

4. A shifting in curriculum weighting for gifted students toward affective, personal and creative goals calls for a corresponding shift in what is to be evaluated.

5. When work in the regular curriculum is adjusted for the gifted student in order to accommodate the special provisions made to meet his/her needs, evaluation procedures must take cognizance of the circumstance.

6. Acceleration of gifted students into advance curriculum levels, usually in one or two subjects, introduces some mismatch between the advanced levels tests and the co-curricular level of student achievement; in most cases the mismatch can be tolerated but should be recognized notwithstanding.

7. Intentions or expectations as specified by objectives can only be regarded as tentative, since it is difficult to establish the worth of any single objective by itself, when the interaction of objectives is unknown.

1. To the degree that students are given Individualized Education Plans (IEP) or Individual Program Plans (IPP), evaluation must also be individualized and the more individualized the evaluation, the more accurate it will likely be.

Evaluation is always individualized to the extent that one receives his/her own grade or that the comments of the teacher are directed to a particular student. However, too often, evaluation of progress is determined relative to what others do with little attention given to what the student intended to accomplish or was expected to accomplish. The IEP and IPP are expected to individualize learning objectives and methods. The extent to which the individual objectives recorded are individually evaluated, will determine the accuracy of the results.

This has implications for the workload of the teacher. If there are few students, it is quite simple to meet with each student individually and develop objectives specific to that student and also to monitor the progress. As the number of students increases, the workload increases and the capacity of the teacher is taxed to the point that the task of evaluation on the prescribed basis becomes all but impossible. This has resulted in teachers asking for a more simplified scheme of monitoring progress such as a checklist.

2. To the extent that objectives are not, or cannot, be defined by objectively decidable criteria, subjective measures must be used.

It is not always easy to define some objectives in ways that are or will be easily measured or graduated. A number of curriculum objectives for gifted students should be built on quality, not quantity. Evaluating student outcomes on these objectives will necessarily demand use of subjective assessment of the data, even though such data are not easy to codify, store, retrieve, or use.

This is especially true for the affective domain. How does one determine if there has been increased awareness or increased interest in a topic? The teacher can ask the student and also make observations of "time on task" and insights about the information as expressed in conversations but these are subjective measures which are not easily quantifiable. While accomplishments of some objectives can be easily determined, there are many that must be left to the professional judgement of the teacher.

3. Where the curriculum is designed to serve a complex of objectives rather than atomic ones, the evaluation must similarly address the complex as an organic whole.

Often in an attempt to measure progress objectively, teachers analyze the goal into minute pieces. The student is then left to work on the small pieces with the expectation that achievement on each will result in reaching the larger goal. While this may work in some cases, it will not always be successful as students get caught up in details and lose sight of the totality. It is also quite possible that some goals cannot be atomized and to attempt the feat is only to destroy something worthwhile. It is better in instances where the curriculum is designed to develop many skills which are not easily discernible from one another, to evaluate them as a group even though there is a seeming loss in accuracy.

4. A shifting in curriculum weighting for gifted students toward affective, personal and creative goals calls for a corresponding shift in what is to be evaluated.

It is important that the objectives of the regular curriculum be achieved for all children but the gifted should also be given an opportunity to go beyond. When these students are allowed to develop other goals and objectives, the evaluation should reflect this shift appropriately.

5. When work in the regular curriculum is adjusted for the gifted student in order to accommodate the special provisions made to meet his/her needs, evaluation procedures must take cognizance of the circumstance.

There are times when a gifted student needs to have the regular curriculum altered to meet his/her special needs and there should be a corresponding alteration in the evaluation criteria.

6. Acceleration of gifted students into advance curriculum levels, usually in one or two subjects, introduces some mismatch between the advanced levels tests and the co-curricular level of student achievement; in most cases the mismatch can be tolerated but should be recognized notwithstanding.

Standardized tests are built on the assumption that students will have some test items too difficult to answer, therefore there is a sufficient spread of scores among test-takers and no student will have a perfect score. When a student does answer all questions correctly, teachers often give the student the next higher level of the test, but unfortunately the test was not normed with this type of student in the sample. Therefore this test is technically not appropriate if the norms are to be used. This mismatch should be recognized by the teacher for the subject(s) in which this occurs.

7. Intentions or expectations as specified by objectives can only be regarded as tentative, since it is difficult to establish the worth of any single objective by itself, when the interaction of objectives is unknown.

While objectives can be thought of at the beginning of any unit, project, or even the school year, these should be subject to modification with added knowledge and experience. In this sense they are tentative and based on expectation. As the project proceeds, the objectives may change, be enlarged, dropped, and new ones developed, or modified in some way as it takes on a unique form under the direction of the individual.

This means that even though the project was done before with another student or under the direction of another teacher, it is still novel. Each person's interests and approach are different and there is always new knowledge and insights added since the last time the project was undertaken. It is also important to realize that as objectives interact, they create a new alchemy which may not in the beginning be foreseen or comprehended but which might produce different and even more worthwhile insights. The visible end products might also need to be altered to match the new objectives. It is important that the gifted child be allowed to start with some objectives in mind but be given the latitude to make changes that seem appropriate. In turn the final evaluation should take into account the changes, the extent to which these have altered the initial direction, and whether the final performance or product is an improvement over what might have occurred without the modifications.

Such judgements are necessary but are usually subjective in nature. They help the child understand that learning is developmental and judgements or evaluations often formative, and that even what is produced is only the beginning of what might be achieved given more time, interest, and further resources.

## ISSUES AND COMMENTS RELATED TO EVALUATING THE GIFTED

Objectives as a basis of evaluation. Evaluation of the gifted should be in relation to their goals or objectives for achievement. This means that more effort needs to be expended in helping these students set appropriate goals of learning and subsequently evaluating them in relation to the goals. This could also apply to the regular curriculum of Alberta Education, i.e. we need to help children determine why the prescribed knowledge is necessary and then help them determine the best way to acquire this knowledge. The student may learn best through independent study and then again learning may be better accomplished through class projects and discussion. There is a place for formal evaluation of content or knowledge but for gifted students the emphasis should be placed upon cognitive manipulation of the knowledge.

---

## ISSUES IN EVALUATING THE GIFTED

---

1. If evaluation should be in relation to previously set goals, objectives and methods, who sets the criteria of performance or achievement?

2. How are the expectations of Alberta Education addressed in a curriculum for the gifted?

3. Should the gifted be compared to the performance of children in the regular classroom, only gifted peers, to their own previous learning and performance, or some combination of these?

4. Do the gifted always show large gains in improvement?

5. If the gifted student fails to find challenge in the program, who should accept responsibility and should the student still be evaluated?

6. How should unique performances, products, ideas, etc., be evaluated?

Teachers should be able to assess the knowledge base of a child in relation to the suggested guidelines set out by Alberta Education. After the child's needs are assessed, the teacher is in a position to combine the required curriculum with challenging curricular ideas developed to meet the special needs and interests of the gifted learner. Once the teacher is satisfied that the child has the necessary basic skills and knowledge, the emphasis should become one of using the knowledge and skills in solving pragmatic problems or in allowing the child freedom to explore a field of interest to acquire greater insights or manipulate concepts to find new relationships that have challenge and intrigue. Evaluation of the gifted child should gradually shift from being teacher-directed to being more self-directed and based upon the goals and objectives set out by the learner. While the child needs to have help in setting goals, and determining how to evaluate performance and achievement based upon the goals, it is the ultimate learning experience of the mature learner to gain insight into setting goals and performing the necessary evaluations independently. It is necessary for teachers of the gifted to keep this objective in mind and promote the shift towards independence in learning and evaluating learning.

Comparison to peers or self-selected goals? Evaluation in classrooms is normally based on a comparison of one child's achievement with that of another, as well as on some concept in the teacher's mind of adequate acquisition of knowledge and skills. With the gifted it becomes less necessary to evaluate a child relative to the performance of others, but rather it becomes necessary to find the absolute criteria against which the performance should be judged. One reason for this is because the gifted child often stands alone at the top of the class with no equal and a comparison to gifted peers may not be possible. The child is then often compared to older children. Is this fair to the child? What should be the expectation?

This is no different than the individual who wants to better his/her athletic performance. There is a certain amount of knowledge necessary about what constitutes ultimate performance and this can be combined with the knowledge about how to improve performance. An individual, knowing the record to be beaten, and knowing his own weaknesses and strengths can then synthesize the knowledge to produce appropriate strategies that will improve performance through a rigorous practice regimen. The performance may then be continually evaluated both by the athlete and the coach, or another specialist, based on the established set of criteria and the objective.

This scheme could apply equally in academic endeavors or it could just as well apply to the visual and performing arts. The need is to first establish the criteria upon which performance is to be judged and then have sufficient expertise on the part of the teacher: a) to help the child determine the necessary knowledge and skills, whether summative or established through a developmental process, b) to help the child assess his/her own knowledge and skills, c) to help the child learn how to improve performance based on his/her present level of development, and (d) to help the child carry on continual self-evaluation.

Improvement Gains. As part of the evaluation process, a teacher must help children learn to discern improvement which might only occur in small amounts and be perceptible to only the most astute observer. The child should also learn to have confidence in the teacher as an expert in the subject field, as an able observer of behavior, and as an interested but impartial evaluator. This places the teacher in the position of the expert and unbiased judge, and is one reason why we need to screen teachers of the gifted. These teachers must have sufficient expertise in a field to give assistance to the child beyond what would normally be expected in the regular curriculum. This may be the reason why some teachers have been able to assist their gifted students while others have been ineffective.

Some students in programs for the gifted, even in special schools for gifted children, have complained that there was insufficient challenge. It could be that there is insufficient challenge because the teachers did not know enough or rather did not know enough about how to teach the children to gain expertise independent of classroom teachers and also how to evaluate their performance. If the teacher is not sufficiently knowledgeable in a given field to provide the opportunity for challenge, or cannot help the child to independently set the challenge and evaluate the performance, then the child is left on his own to provide a challenge. Too often boredom and dissatisfaction are the result.

Evaluation for adults. Evaluation in the adult world is frequently based upon a comparison with the performance of others. In fields where something new is created, evaluation is based upon a different set of criteria, usually the adequacy of the solution to a particular problem. It may be the cure for a disease, the best method of lifting a load from point A to point B or it might be a unique conceptualization of various concepts relative to one another. It seems that this is why we have begun to view giftedness, not so much as how children perform, but as what questions they pose. The questions posed suggest the quality of cognitive activity and the relevance of conceptual relationships. This question might ultimately lead to new solutions and creations.

Being able to judge or evaluate one's own performance or ideas against some criterion, or set of criteria, thus becomes one of the most useful tools of learning. This may say that we need more teachers who can help children pose questions than teachers who can help children find solutions to mundane problems. While teachers of the gifted may not need to find solutions to esoteric problems, they should have adequate knowledge in the field of study undertaken by the child. Of most importance, however, is the teacher's ability to help children pose questions and manipulate ideas which will in turn bring about insightful understanding. This is not to downplay the fact that insight can be gained at each point on the continuum of learning, depending upon where the individual is stationed, but rather suggests that the teacher needs to be so attuned to the child that he/she can sense what questions and challenges are needed to spur development at any particular point in time.

## DEVELOPING EVALUATION CRITERIA

The criteria upon which to evaluate a project, presentation or any bit of learning can be of benefit to the learner as he/she anticipates a predetermined end point. There are many questions to be answered by the teacher, by the student, by peers and by others who are interested in the progress of the learner. Not all of the objectives of each endeavor are discernible at the beginning nor should they be. If the individual is aware that many benefits can accrue as the project progresses and that some objectives will need to be modified or even dropped, this too is worthwhile learning. Each project can be a new experience, with new insights that can have a marked effect upon the learner.

In many classrooms providing enrichment and special activities for the gifted, there is an emphasis on projects and presentations. There should be an understanding on the part of the learner and the teacher as to why projects are a part of the curriculum and in what way they are related to learning.

Some aids to learning. It must be recognized that learning has many facets and that prerequisite skills and knowledge are essential. How previous knowledge and experience is related to new learning and the extent to which quantity and quality are germane is not clearly understood. Suffice it to say that the gifted child seems to come with far more of the prerequisites in place than we would normally expect in the regular classroom. No one seems to be quite sure how this occurs but teachers can attest to the fact that it is a real phenomenon.

Educators are also aware that there are several steps in the learning process and that feedback on progress as a part of the learning environment is absolutely necessary. It has been found that where the individual understands the reason for learning, there is more meaningfulness and hence learning is quicker, more comprehensive and lasts longer. This is readily applicable to the gifted learner because he/she seems able to understand easily the reason for learning and can find problems that are worthwhile exploring either on his/her own or with minimal guidance from others. Once the area has been defined then information must be obtained either through direct instruction or by some other method whereby the learner takes an active part. Active learning involves the use of cognitive skills which organize information, relate it to past experiences and learning and enable the learner to make applications, judgements and forecasts about its possible use. The learner usually needs practice under guidance with adequate feedback on accuracy and quality of achievement. After sufficient guided practice the learner is ready to attempt practice independent of a guide or teacher but still requires feedback in some form on accuracy. This feedback may come from the learner, from peers or from some other evaluation source. It is important to realize that evaluations of achievements or attempts to achieve are a part of life and an essential aspect of learning about living.

Another aspect of learning that is sometimes overlooked is the communication of learning to an audience. Projects are often devised without adequate thought about the best method of acquiring information, given the limitations of the environment, or how the information could be reported to others. The information assembled to address a problem is often evaluated with little thought given to the cognitive manipulation of the information required to address the questions that might be raised and the implications of proposed answers.

Why is all of this information necessary to keep in mind as one considers evaluation? Because as one begins to assist the gifted child to consider how projects and performances could be assessed, all appropriate and relevant factors should be discussed.

The following categories of information and attending details may be useful as teachers develop evaluation instruments for themselves, assist students to conduct self-evaluations, and endeavor to develop instruments with the aid of gifted children in peer evaluation situations.

## REPORTING STUDENT EVALUATION

Teachers were asked about their procedures for reporting student evaluation. There seems to be a need to address some of the concerns. Some of the more common methods of giving feedback and reporting are listed below.

1. Evaluation is provided to students in the form of report cards in some instances. Often there is no grade given for work done in enrichment classes but there might be a comment on the report card about progress. Too often there is no mention of the enrichment session at all.

2. In most cases gifted students are given feedback on their work in an oral fashion and the information is recorded by the teacher in the form of comments as an anecdotal record about the projects, performances, etc. and this is reported to parents in teacher-parent interviews.

3. In some cases students and teachers have a conference during which the finished project is discussed and the student is able to provide some insight as to why the work was done in a particular way. Such conferences seem to have greater merit than teacher comments on a written project or verbal feedback only.

4. When examinations are used as a part of the evaluation process, students receive a mark and usually know where they are in relation to other members of the class. These marks are usually transformed into a grade that is reported to both students and parents.

## SPECIFIC SUGGESTIONS FOR EVALUATING GIFTED STUDENTS

1. Use normal pencil and paper tests constructed by the classroom teacher to measure the skills and knowledge of

the regular curriculum. These tests should at least evidence face validity, i.e. the questions should appear to measure the knowledge under consideration. These tests should attempt to evaluate the students' ability to operate with the knowledge using several cognitive levels from memory to problem solving.

2. Avoid the use of standardized tests for evaluating students' performance in areas specific to the gifted. Such tests do not adequately discriminate among the gifted students because they lack validity. They seldom measure what needs to measured.

3. When projects, performances and reports are used to measure the student's ability, there should be an attempt at the beginning to decide what objectives are to be achieved. Is the teacher looking for commitment to the task? Is there supposed to be a transfer of knowledge from one area to another? Is neatness a factor? What about verbal expression and clarity, including grammar, spelling, etc.? What about thought processes such as drawing inferences and conclusions, analysis, etc.?

   With the objectives clearly outlined beforehand and a discussion of the possible alternative methods of achieving the intended results, the student is allowed some freedom of choice as to the best method to employ. It is not only necessary for the student to select the method of learning but also the method of communicating the learning to an audience composed of people who are in effect both learners and observers. Once the student makes a choice, he/she could be evaluated on such aspects as completeness of background knowledge, application of the knowledge, clarity of presentation, how close the performance is to a perfect performance based on known criteria, new insights gained through the process, or on such other criteria deemed appropriate since these become the objectives for achievement.

   The evaluation should not be totally performed by the instructor but there should be cooperation with the student so that self-evaluation becomes a part of the total process. Self-evaluation should be accorded more weight concomitant with increased experience with a particular topic or skill.

4. Peer evaluation is also a vital part of the evaluation process and is a necessary part of learning to be an independent learner. It is worthwhile to solicit evaluations from student peers but these students should

know the criteria upon which evaluation is being based or should have an active role in developing the criteria for evaluation.

It is important to let students be involved in evaluation because it benefits all concerned. The student learns that just as feedback is necessary for improving performance, self-evaluation is necessary to improve our performance when no one else is present and can give a sense of independence in learning and achieving. The student being evaluated learns that criticism of peers can also provide important and valuable feedback in improving any performance. Students learn the basis for evaluation and begin to appreciate the need for criteria on which to make sound judgments. Students begin to learn the difference between subjective and objective judgmental criteria and the role of each in the evaluation process. Teachers find out how important evaluation can be and also how much students can contribute both through peer evaluation and self-evaluation.

5. It is important to seek the opinions of other teachers in regular classes and those who are teaching other gifted children. These comparisons help to keep perspective and provide an opportunity for colleagues to assist by bringing their experience to bear.

6. There is a place for the use of professionals and specialists in an area to evaluate performance. We often invite specialists to adjudicate musical talent festivals for this purpose. The festival provides an appreciative audience for children and the adjudicator gives feedback which can act as a challenge for improvement in this safe environment.

7. Use of anecdotal records is common in evaluating the gifted. This requires that the teacher become an astute observer of relevant behavior. Teachers of the gifted claim this works well when the numbers are few but becomes less accurate and the paper work burdensome with large numbers. In this case it is suggested that checklists developed by the teacher which take into account the objectives to be achieved should be used. These can be jointly developed with a particular student and/or with a class.

8. Shared evaluation between teacher and student has proven to be popular with gifted students. This needs explaining since it goes beyond cooperative evaluation.

This means that the teacher evaluates the student on his/her achievements and the student in turn evaluates the teacher on the way in which he/she has provided guidance and has established the learning environment. In this type of evaluation the teacher becomes much more aware of his/her role as a partner in the learning process and there is a greater understanding on the part of both about the objectives to be achieved and the methods to be employed in learning and reporting on the learning.

9. Teachers use discussion as one of their prime teaching methods but have expressed concern about how this can be adequately evaluated. When asked what was gained from discussions, the following ideas were given: a) children learn that others think as they do, b) children find out how to approach situations and think, c) children gain an interpersonal understanding of skills, d) children are able to discuss their problems with caring teachers and peers who understand how they think and who can give honest feedback, e) students are helped to sort out what they see as a very complex world because they concentrate and focus on a topic in greater depth, f) because some children think in different forms, i.e. one student thinks in color and one in shapes, they help others understand their thinking processes and thus become more accepted, and g) affective skills are developed through discussions - cannot divorce the affective domain from the cognitive in learning.

   Evaluating discussions is difficult. The teacher must relate to each child based on ideas, conclusions, thought processes, responses to feelings, and many other things which cannot be easily quantified. It is no wonder that sometimes the diversity of opinion, insight, feelings and knowledge leaves the teacher feeling inadequate and distraught with the concept of evaluation. There is no easy way to evaluate discussions but if teachers believe this is necessary they should carefully develop the criteria in advance and then discuss it further with the children to receive their input.

10. There is great concern on the part of all teachers about how to evaluate the affective domain or whether it should even be attempted. Some suggestions about how this might be approached were provided by the teachers involved in this study. a) Use self-descriptions before and after experiences, or even at the beginning and end of the year for a method of self-evaluation, b) Note the degree of risk-taking as a measure of increased

self-confidence, c) Note the physical demeanor; i.e. body language tells something about self-confidence such as the child who stands with head erect instead of down, no longer acts out, change in voice tone, etc., d) Note peer relations, i.e. less isolation, less competitiveness, not threatened by peers, etc., e) Note attitude change such as the excitement about activities. This needs to be sampled over a long period of time, f) Note how a child creates his/her own audience when they have self-confidence, and g) Note the reaction to the teacher, such as more accepting and loose as the child gains in self-confidence.

Teachers have found that one of the best devices for noting change in the affective area is through the use of journals where children can feel comfortable in expressing attitudes, feelings, and learning. Of course the testimonials of peers, other teachers, and parents can be beneficial in this evaluation also.

## SUGGESTIONS FOR EVALUATING GIFTED STUDENTS

1. Use normal paper and pencil tests constructed by the classroom teacher to measure skills and knowledge of the regular curriculum.
2. Avoid use of standardized tests for evaluating performance in areas specific to the curriculum for the gifted.
3. When projects, performances and reports are used to measure achievement, the main objectives should be determined at the beginning.
4. Peer evaluation should be used as a part of the process but they need help to develop the criteria.
5. Opinions of other teachers who have the child in their classes should be requested.
6. Professionals and specialists may be used to evaluate special performances and projects.
7. Anecdotal recordings of relevant behavior may appropriately be used as evaluation data.

8. Shared evaluation, i.e. teacher evaluates the student and the student evaluates the teacher on appropriate criteria in the teaching- learning environment.

9. Evaluating discussions usually requires subjective judgement but the objectives to be evaluated should be analyzed with caution.

10. Evaluation of the affective gains might be attempted through: a) use of self- descriptions, b) degree of increased risk- taking, c) change in physical demeanor, d) change in peer relations, e) attitude change, f) creation of own audience, g) change in acceptance of, and relation to, the teacher.

---

## FOSTERING SELF EVALUATION

It is a belief that students should learn to be more self reliant in terms of their own educational pursuits. The gifted child already seems to be pointed in this direction. While the principle of independence in learning, and self evaluation to aid that learning may be sound, there are many aspects which must be addressed. The following is an attempt to isolate some of the pieces that must receive attention for helping students conduct a thorough self evaluation. It is recognized that self-evaluation will be a developmental process and that some items are more important than others and that some have greater applicability to the classroom setting than do others. Nevertheless as one begins to concentrate on the obvious concerns, it prompts others to appear. It is hoped that the concerns raised here will assist teachers and students as they attend to the evaluation of achievement.

An assumption is being made that learning goes on constantly within the individual but it is not recognized by others, whether it be peers or teachers, until a response that can be observed is communicated. It is also assumed that feedback is a necessary part of the learning process and that once an individual makes an overt response, the comments of others who form an audience, provide the learner with feedback which can be used to improve performance.

Unfortunately the student suffers if the products or performances are poorly presented; have deficits in style, format, content, or organization; are untidy, or in some other way fail to meet the expectations of the audience. The child may have learned but because he/she could not communicate the learning, a judgement may be rendered that seems unfair. This means that teachers must not only help children determine the objectives of learning but help these children display learning for those who are in a position to provide feedback and make judgements about the worthwhileness of the products and performances. This argument also leads to the suggestion that students be given an opportunity to provide an evaluation of their learning which could be incorporated into the final judgement.

## Considerations for Self Evaluation

It seems possible that if the students learn appropriate criteria for evaluating their performance, they will in turn routinely establish these into the learning style. This does not mean that the criteria cannot be updated with developed skills and knowledge or that performance will not continue to improve. It merely suggests that if the student knows the target and is given adequate information about the accuracy of each performance in relation to it, learning will proceed at a faster rate. Through being informed the student is also likely to be happier about the circumstance.

The following list of criteria for judging a product or performance is provided as a beginning point to help students with evaluating themselves. The list has been developed from the perspective of the learner but it could also be adapted by outside observers.

### Criteria

1. Project Initiation
   -was the idea my own or someone elses?
   -did I get enough help in the planning? why/why not?

2. Learning Objectives
   -were these accomplished?
   -are there some not finished? why?
   -can you suggest methods for achievement?
   -were all aspects of the problem addressed?
   -what new objectives developed during the project?

3. Procedure (Method)
   -was the task easy/difficult? Why?
   -did you gain interest or lose interest during the project? Why?
   -did you have enough knowledge to begin the project?
   -what advice could you give to others attempting this type of project

4. Product or Performance
   - organized?
   - complete/thorough treatment?
   - neat?
   - aesthetically pleasing?
   - unusual/creative?
   - grammatically correct?
   - for oral presentation one might consider voice quality, and demeanor such as eye contact with audience, appeal, etc.
   - did I communicate well with the audience? How do I know?
   - what would have made my product/performance more appealing?

5. Individual Development
   - knowledge
     - ability to locate information?
     - organization of the knowledge?
     - efficiency?
     - unexpected gains?
     - new insights?
     - things to avoid?
     - importance? in what way(s)?

   - maturity (understanding of self)
     - understanding of my own interests? work habits? abilities?
     - greater independence?
     - interactions with others?
     - effect of the project on me? Why?

   - satisfaction
     - am I satisfied? Why/why not?
     - was the project/performance worthwhile? Why/why not?
     - would I do this project again?
     - how would I change the approach?

   - improvement
     - how have I improved from my last project?
     - how did I incorporate previous learning into this project?

## Considerations for Peer Evaluation

We all need information from our peers about how our products and performances affect others. Too often however, peer evaluations are not well established in appropriate criteria which have been thoroughly analyzed and tested objectively. There is the tendency to evaluate on the basis of a superficial coverage and from the affect upon emotions. It seems that if students were taught the basis for evaluation, feedback would be more meaningful to their peers and they would also have gained some valuable insights into evaluation.

The following list of considerations is far from complete. It is an attempt to help teachers convey a message to students evaluating their peers, that appropriate criteria must be established, understood, and that evaluation is a highly significant cognitive process. During the process the evaluator should develop sound reasons for judgements and be able to assess whether the reasons are due to emotional effects.

Criteria

1. Criteria for Evaluation
   - were these provided?
   - how were they developed? basis?
   - are they appropriate?
2. Project
   - worthwhileness? why?
   - was it challenging to your peers? why/why not?
3. Learning Objectives
   - were they clearly stated? could you tell what the intent was from the presentation?
   - were there unstated assumptions that affected the product/performance?
4. Procedure (Method)
   - was the procedure for accomplishing the task appropriate? Why/why not?
   - what other methods might have been used?
5. Product/Performance
   - was the planning and preparation adequate? Why/why not?
   - was it well organized?
   - was the treatment thorough or complete? detailed?
   - were all aspects of the problem addressed? which ones were not?
   - was the product/performance appealing? Why/why not?
   - was the product/performance unusual/creative
   - did the product/performance show attention to detail?
   - was the audience adequately prepared in advance for

the product or performance? how?
- -was the project/performance persuasive? Why/why not?
- -were there mannerisms/attributes of the product or performance that were distracting?
- -was there practical value? What?
- -were conclusions appropriate? Why?
- -what suggestions do you have for improvement? are these meaningful? to whom?

6. Effect on Audience
   - -how do you feel about the product/performance? (e.g. enlightened, interested, bored, embarrassed, entertained, etc.)
   - -has your opinion of the presenter been changed? how?
   - -would you seek advice on the topic from the presenter? why?
   - -were you satisfied? why/why not?

7. Evaluation
   - -did I give full attention to the product/performance in order to give a thorough evaluation?
   - -have I been fair and objective? Why/why not?
   - -what about the product/performance influences me most? why?

## SUMMARY STATEMENT

The evaluation of students in special provisions for the gifted is not an easy task. It is fraught with difficulty and inequity. Yet students want feedback about their achievement and they do want to improve their performance. Through a study of the issues related to evaluating the gifted, some principles and criteria have been developed which will hopefully help teachers with the evaluation process.

# CHAPTER FIVE

# OBSERVATIONS AND IMPLICATIONS FOR EVALUATION

## OBSERVATIONS ON EVALUATION STRATEGIES

During the course of the project, some observations about using the evaluation strategy, and some implications for evaluating student growth and programs became obvious. They are offered in this chapter in the expectation that they might be useful and not obvious to all of the readers.

An interesting feature of the strategy used, is that it not only permits the recognition and use of known models in its various procedures, but even if the models are unfamiliar to the evaluator, they will still function as a part of the strategy. They are inherent, not an addition. In illustration, three kinds of models are often encountered in program evaluation. These are the connoisseurship models, the discrepancy models, and the adversary models.

Connoisseurship models rely on a "thick description", i.e. a very detailed description of the history, organization, curriculum component and the many interactions within the district in order for the reader to gain a comprehensive view as if present during program development and thus to understand what is happening and why. Evaluators are then able to make proper assessments and recommendations. The question, "Are the needs of the students identified being met in the provisions made for them?", normally invokes connoisseurship model procedures. Needs are identified, evidence of program benefits assembled, and then informed professional judgment is required to relate the two and reach some decision about adequacy. The strategy used in this project incorporates this aspect even though the evaluator may not be familiar with the connoisseurship model.

The discrepancy model works with such questions as, "Is the system approved process of identification used throughout the district?". The expected procedures are known and what remains is to gather data on procedures and routines actually in use and then make the necessary comparisons. This process will also occur in the strategy whether the model is consciously recognized or not.

Adversary models follow a debate or legal scheme. Both sides of the case are presented and a decision made by a jury-like panel or a judge. If the purpose of the project had been summative in nature, i.e. to decide on whether to keep or drop the program, the typical routine of this model would have been to articulate the questions, determine data to be used, and indicate procedures appropriate to the decision making processes. This too is incorporated into the suggested strategy.

Formative evaluation tends to rely on the first two models (connoisseurship and discrepancy) while summative tends to be more adversarial. However, throughout the various stages of the recommended strategy, all necessary aspects of the other models are adapted. The level of attention required from experts in the field is thus identified, and the proper knowledge applied and incorporated into the appropriate judgemental component.

## Implementing the Strategy

Care must be taken from the beginning to avoid overwhelming the evaluator with irrelevant or unimportant background data, yet there must be enough so that the evaluator becomes familiar with the system. If the system is large, a longer time period and more data will be needed to inform the evaluator(s). If, as sometimes happens, it becomes necessary to use a panel of judges or experts in the field to interpret data and render judgements, the amount of data needed may be increased and the additional load must be tolerated. But taking care to gather only the data needed is one way of making report building a manageable task.

Related to both the reporting and the accuracy of the data is the method of gathering it. Several comments can be made on the value of the interview method for this purpose. It can be a good means of validating data and data gathering processes. It saves the time of the person interviewed and allows the interviewer to be more certain of the accuracy and usefulness of what is gathered.

Flexibility of evaluation strategy in this project was necessary because varying sizes of school systems were to be evaluated, as well as various sizes of the evaluation projects themselves. The strategy functioned well over the variations in both dimensions as hoped.

One caution should be reiterated. There is a tendency to rush and downsize the pilot phase in order to get on with the job. But the quality of work in the pilot phase in great measure controls the quality of the evaluation results and the ease with which they are attained.

Individual Program Plans (IPP's) and related (although not usually synonymous) Individual Education Plans (IEP's), Individual Learning Plans (ILP's), etc. suggested a way of evaluating program outcomes, i.e. student development. They are being used in the education and evaluation of the mentally handicapped and it was hoped that they could become almost immediately functional in evaluating the gifted. This hope is not realizable in the near future. There are two impediments. First, attention is just now turning to the use of IPP's in planning educational experiences for the gifted, and so even when they are used, their form and content does not easily adapt to their use in evaluating student growth. Second, teachers of gifted students find them too time consuming since curriculum concerns preempt most of the limited time available. Development work is being done, however, and the hope may still be realized. Strathcona County, for example, has developed and will pilot, in September 1986, a form combining information normally found on cumulative records, such as personal information, a summary of the student's academic record and a list of items for teacher rating in the following areas: Motivation, Social and Affective factors, Communication, Critical Thinking, Creativity, and Research and Organization. (See Appendix B ) The form also makes provision for noting student strengths, input from the homeroom teacher, and specific plans and recommendations for the student's individual education plan. Regrettably, however, it must be reported that the mainstream education-for-the-gifted enterprise is not ready to use IPP's for evaluating student development, although the promise is still alive.

In conclusion, an interesting observation can be made about the synergism of the strategy itself. It was found that each step seemed to illumine and point the way, as well as interact with the next and the other steps. The process was interactive and dynamic.

## Implications

The impact on a school system's evaluation services from adding an education program for gifted students goes substantially beyond the effect of adding just another new program. Programs for the gifted have fundamental differences which create unusual and extensive evaluation needs and demands, much like those for slow learners. Some can be met by existing services, others go beyond them, and yet still others suggest directions for research and development. In student development evaluation, new kinds of data sources must be considered and an increased focus is placed on objectives of both instruction and learning. In program evaluation, there are new components to be evaluated and new functions demanded of some of the usual components. The majority of the implications, however, grow out of student development evaluation.

## OBSERVATIONS FROM STUDENT EVALUATION

Goals and objectives which determine the learning experiences for gifted students are the source of the most complicated tasks in evaluating student outcomes. The goal, (see Alberta Education, "Goals of Education") universal education in general and for the gifted in particular, is to enable them to "realize their potential to self and to society". This goal reaches so far into the future that the chains and sequences of objectives which lead to it cannot be securely known. The individualized goal, moreover, is not the same for all gifted children but varies from person to person and hence elaborate chains of objectives must also vary from person to person.

The primary task of a system's evaluation services, is to assist in the achievement of the educational goals. One important means of doing this is by providing the necessary feedback information about the extent to which objectives are being achieved, thus serving to keep the program on target. The need for such service is beginning to be recognized. One of the school systems in the project has even instituted a long range evaluation program to identify the effects of special programs for the gifted on their educational careers. This is a long term project requiring continuing attention. Extension of such a program to determine the effects of special provisions on work careers would provide important additional information.

The second, and more important task of evaluation services is to improve the effectiveness of the objectives themselves. By evaluating and providing feedback on the effectiveness of objectives as a means to achieve the long range goals, and also as a means to define short range expectations for student growth, evaluation services will be able to provide a major contribution to the formative evaluation, i.e. the development of programs for gifted students. This kind of feedback will focus directly on the most important part of the education provisions, the teaching-learning experiences, thus giving increasingly informed direction toward efforts to achieve long range goals. Meanwhile the evaluation efforts will serve to provide technical help in defining short range expectations and the criteria which signify various achievement levels. Neither task, however, is routine or easy.

Building toward future goals will involve screening suitable objectives from the many curriculum offerings available for gifted students. The processes employed are likely to shift from the computational toward the judgemental with some corresponding changes in the role of evaluators. Assessing the wisdom of the decisions which result from these processes will require careful long range record keeping. The results will repay the effort by providing feedback to improve both processes and judgements. The judgemental processes themselves may develop in unfamiliar patterns probably involving parent, student, and professional experts to an increasing extent.

Teachers of gifted students identify some expectations by various processes (often IPP's) which frequently involve student(s) and others. The expectations are usually attached to some student outcome: a product or project. The art of setting expectations which challenge the learner appropriately, may be part of the 'craft' of teaching and so left to teachers, but they need help in determining how, and to what extent, meeting the expectations determines achievement of objectives. This is an evaluation procedure in which objectives are used in defining expectations so that they may then be considered for their usefulness in achieving long term goals. IPP's may prove to be an effective tool, but the research ground work necessary for their development has still to be laid .

There are, at the present time at least, portions of the endeavor toward educating the gifted, where expectations are almost completely unknown. Approaches to evaluation of outcomes in such situations exist but this is a new game for teachers and perhaps some evaluators. This is another non-routine task for an evaluator.

## OBSERVATIONS FROM PROGRAM EVALUATION

Definition of who is to be served and identification are two important components of programs for the gifted. They have no close counterpart in regular education programs. The obvious requirement is that an evaluator have some expertise and background knowledge of what purposes these two components serve and how they articulate with the total program.

Special provisions activities in programs for the gifted present much more important and extensive demands on evaluation services than do regular programs. The previous discussion on student evaluation points up the fact that the objectives for gifted student outcomes are not yet totally established. The uncertainty about objectives implies uncertainty about curriculum. This is not surprising because recognition of the inappropriateness of regular curricula for gifted students was the point of departure for building suitable programs for them. The effects of the departure, however, are extensive. The new curricula are the result of some combination of curriculum development, adaptation, adoption, and selection. Hence evaluation of programs for the gifted must consider the basis of curriculum building, the effectiveness and appropriateness of the curriculum, as well as the worth of the product. These are substantial additions to the evaluation services required in regular programs where the worth of the curriculum is usually taken for granted and very little curriculum development takes place at the teacher level.

The nature of programs for gifted students imposes another demand. It can be concluded as a result of this project that no one existing evaluation model completely meets the evaluation needs of programs for the gifted. Consequently it is necessary either to piece together a new model using parts from existing models or develop a new approach. In either case, dealing with new kinds of data and making evaluative judgements using different processes is involved. These new requirements have implications for evaluation services and call for additional support services.

### Implications

The following are suggestions in order to evaluate properly special provisions for gifted students.

1. There needs to be a permanency or continuance to the service in order to provide formative feedback to direct a program toward the long term goals for students. It

will also facilitate formative assessment of instruments and procedures for student development evaluation and program revision.

2. In-depth understanding of programs for gifted students is needed to ensure that all of the unusual features of program and student evaluations, as noted in the two preceding sections, receive appropriate attention.

3. Facilities for keeping current with developments in education for the gifted are needed to maintain quality feedback.

4. Communication structures are needed to ensure adequate feedback to shape program design, revise programs, improve teaching-learning activities, and to influence teacher in-service.

5. Capability (time, support staff, etc.) of working with teachers, parents, and students in developing and/or testing processes and instruments useful in student evaluation is needed.

# APPENDIX A

## LITERATURE SURVEY

The consensus of opinion is that current practices in evaluating education for gifted students need improvement. In a 1982 report, based on a national survey to the advisory panel, U.S. Office of Gifted and Talented, Weiss and Gallagher state,

> "The final conclusion that the reviewers reached in this volume is that there is a serious need for a systematic and organized effort to improve the design of program evaluation efforts in the schools, if such evaluation is to be considered a valuable tool for future program decision making", (p.3.).

These authors did not change this opinion during the course of the survey,

> "The final conclusion of the authors is that the area of gifted education remains a fertile, but still largely unexplored, field in need of consistent and much more numerous formative and summative evaluations", (Executive Summary).

A similar opinion was reached, independently, as a result of a more intensive, albeit less extensive, survey of programs in three provinces and eight states. Sillito and Wilde (1983) report that,

> "... one of the first and most significant conclusions emerging from this survey ... is that

> ... evaluation is a crucial yet neglected area whose quality is in need of upgrading", (p.220).

Archambault (1983) noted that, "The evaluation of programs for the gifted and talented has become an increasingly frustrating challenge for project directors and evaluators ..." (p.12) and that, "by comparison [with the academically disadvantaged], negligible amounts of money have been expended for evaluation ... in the area of gifted education", (p. 13).

Almost a decade ago, Renzulli (1975) advised caution in using any one model of program evaluation. Before exploring five different models by diverse experts including himself (Stake, Stufflebeam, Provus, Eash, Renzulli and Ward) he offered a caveat,

> "A would-be evaluator could easily drown in a sea of complexity if he selected a single model and slavishly tried to stick to it . . . it is probably true that no single model will serve all of the evaluation needs of a given program", (p. 17).

After discussing each of the five models he reiterates the warning,". . . no single model or approach is sufficient for solving all the problems that are likely to occur . . ."

Measurement of program impact, i.e. effectiveness or student related outcomes, is reported as being less than satisfactory. Alvino, McDonnel and Richert (1981), on the basis of a national survey on instruments and practices used in identification, report that,

> "The trends . . . indicate some disturbing recurrent practices . . . From a psychometric point of view, many tests/instruments are being used for purposes and populations completely antithetical to those for which they were intended and were designed", (p. 128).

These authors indicate a misuse and/or a lack of suitable instruments, ". . . rampant use of informal and subjective measures is taking place, which may indicate a lack of suitable instrumentation . . .", (p. 130). That the situation is as bad in evaluation is a reasonable inference. It is, in any case, supported by others.

In the report already noted, Sillito and Wilde indicate that instrument validity is a pervasive problem,

> "The basic assumption . . . which is so often . . . unwarranted, is that the test curriculum adequately fits the actual curriculum presented to the student. In the case of a 'qualitatively different' curriculum, which is a major objective in education for the gifted, the probability of a massive misfit approaches 100 percent!"

Archambault (1983), in discussing the possible use of out-of-level tests to avoid the ceiling effect for gifted students, observed,

> "The first issue here concerns the instructional relevancy of the test . . . one must be assured in selecting [a] . . . test that there is really a match between the content of the test and the instruction delivered through the program", (p. 18).

and advises "avoid out-of-level-testing where possible", (p. 19.

Renzulli (1983) points out that a push for "hard data" is equivalent to mounting pressure to measure "molecular" behaviors and may prove to be both antithetical and destructive of efforts with gifted students to encourage complex "global" behaviors which are necessary for the complex products that the gifted create - such as poems, essays, art, and research. Renzulli insists that, "it is better to have imprecise answers to the right questions rather than precise answers to the wrong questions" and that "it becomes more crucial that we begin to live with more subjectivity in our evaluation efforts", (p. 6). Renzulli opposes excessive influence from the behavioral objectives movement in educating gifted students.

> "Creating behavioral objectives will never meet the diverse requirements of programming for gifted and talented youngsters. The task . . . is far more complex and requires being able to live with some subjectivity", (p. 10).

The apparent point which is being made by Renzulli is that the behavioral objectives approach to evaluation is inherently and irremediably invalid when applied to the complex and higher order skills and behaviors which are the core of a qualitatively different program for gifted students.

Callahan (1983) states that, "the instrumentation which has been used to evaluate programs for the gifted has often been invalid, unreliable, or simply unrelated." Use of standardized achievement tests, she notes, is generally invalid because: a) they address the objectives of the regular curriculum rather than the differentiated curriculum, b) they assume that all gifted students are studying the same content, c) they emphasize mainly skills which are often not part of the special program, and d) their use introduces the regression-toward-the-mean bias. This bias is particularly exacerbated for the gifted. Other validity problems noted by Callahan relate to the unusual heterogeneity among gifted students and their need for individualized programs. She concludes that there are fundamental difficulties, arising in part from instrument problems, with both the experimental/control group and the behavioral objectives designs.

> "There are important issues or problems which have kept both theorists and practitioners from making reasonable judgements about the effectiveness of programs, and thus merit creative efforts to resolve them."

The above researchers (and others) appear to be critical of current practice in evaluating program/student outcomes in education for the gifted, for one or more of the following general reasons:

1. Evaluation is often lacking.

2. Where program evaluation exists, it is too often an add-on and inadequate in design and/or perfunctory in application.
3. Program Evaluation designs in use may be inappropriate.
4. Instrumentation for assessing impact, i.e. effectiveness or student outcomes is scarce and what there is is often either mis-applied or mis-interpreted.
5. The state of the art in impact assessment in education for the gifted is an inadequate basis for program evaluation.

Whether or not the above and similar criticisms are justified, is difficult to judge but there appears to be a dearth of defenders of current practices. An examination of what specifically is thought to be amiss in evaluation practice seems, therefore, to be in order.

## Deficiencies in Current Practice

The targets of the critics range from general philosophies and concepts of design and strategies to the specifics of practice.

Renzulli (1983) expressed an interesting position which merits consideration,

> "I personally do not think that evaluation should be THE major reason for funding or not funding a project . . . I . . . believe that gifted programs need to be supported on philosophical or even legal and moral grounds - not simply financial concerns", (p. 8; author's emphasis).

He would not, however do away with evaluation. "Program personnel have a right to make their programs better as time goes on." His preference would be to emphasize formative evaluation,

> "Good formative evaluation would go right along with the natural sequence and flow of events, and primarily serve to both document and plan a given program's direction and status", (p. 8).

His opposition to the behavioral objectives approach, noted above, appears to be based on the conviction that their use is "reductionist", leading to increasingly "molecular" behavior which neither accurately specifies appropriate goals in educating gifted students nor serves as a basis for establishing their achievement.

Callahan (1983) expresses a complementary position about behavioral objectives. If the objectives are general enough for program staff to use in planning, they are too general for effective use in evaluation. She cites as examples the general statement, "The students will become more creative." and a more specific one, ". . . 90% of the students will achieve a gain of five points in fluency scores on the Torrance Tests of Creative Thinking." According to Callahan,

> "Neither of these positions serves gifted children very well . . . the first, more general statement leaves the evaluation of changes in creative behavior to the whim of the evaluator; it is dependent on the definition of creativity chosen by the evaluator - not the program staff. The second, more technically correct behavioral objective (according to Mager, 1962) is also inappropriate for the following reasons:
> 1. It adopts an extraordinarily narrow view of creativity (or other concepts normally included) . . . in programs for the gifted.
> 2. It implies . . . [a] standard . . . of achievement . . . expected from gifted children . . . a standard that does not exist.
> 3. . . . it assumes that the program will be the same for all children and have the same effects on all children."

Unlike Renzulli, however, Callahan assumes the dilemma is resolvable.

> "It is the responsibility of program developers to define goals as behaviorally as possible. They must insist on the direct measurement of the behaviors defined, and work with evaluators until the instruments selected or constructed do measure the appropriate skills, learning, and achievements", (p. 4).

Callahan does not indicate how this resolution of the dilemma may be achieved nor does she show that it is even possible. There is some likelihood that the different positions taken by Callahan and Renzulli are based on differing concepts of curriculum or even definitions of giftedness.

Sillito and Wilde (1983 p. 13-15) raise the question of which exceptional ability or set of abilities in combination constitute giftedness. They observe,

> "Broader programs which typically include cognitive, creative, high achievement, and visual and/or performing arts capabilities do not lend themselves . . . to tightly planned curricula or to the study of the effects of the various combinations. Eligibility [for admission to the program] is usually based on exceptionality in one or more of the capabilities so that the students served exhibit a . . . variety in combinations of exceptionalities. It is therefore difficult to program and evaluate with any degree of specificity", (p. 15).

They note that a frequent strategy in school systems using the broader definitions is to individualize student programs. The inference is that a priori specification of behavioral outcomes in adequately complete form for groups of gifted students is not feasible.

A different term, "measurable objectives", is used by Hamilton (1981). This term is not necessarily the same as "behavioral objectives" in the highly technical sense made popular in the education community by Robert F. Mager. Creativity can be recognized by many behaviors and products, as can the higher level thought processes, without prespecifying the specific behavioral indicators and the criteria for acceptable performance on each. Hamilton, however, appears to be using the terms almost synonymously. He references Mager and states,

> "Dealing with the outcome question . . . requires statement of some objectives. (p. 545) . . . If your objectives are stated in behavioral terms, criteria for judging whether the outcomes are satisfactory are included and you have only to

> demonstrate that those criteria have been met to establish that your program was effective", (p. 548).

Hamilton was discussing the evaluation of innovative programs, although not specifically for educating gifted students.

In addressing the question of evaluation design, Archambault (1983) states, ". . . the major concern is the inappropriateness of the more traditional designs and the resultant need for alternative designs and procedures." (p. 12) He accepts two concepts: impact and comparison. Impact is the extent to which a program causes change in students. Comparison poses the question of whether alternative programs would have greater impact. Two general categories of evaluation design are cited, the experimental, and the quasi-experimental. Both measure impact and allow for comparison. The first is not feasible and consequently the weaker quasi-experimental design is the other remaining choice.

> "Unfortunately, most programs for the gifted and talented do not allow for the random assignment of participants to groups . . . Thus, the possibility of using the true experimental design is removed. . . . Because the true designs (i.e. the Pre-Post Experimental Design, and The Solomon Four Group Design and the Posttest only Experimental Design) provide the strongest evidence of program impact, evaluation of gifted programs most frequently employ what are regarded as weaker designs", (p. 20,21).

Among the weaker designs is the Non - Equivalent Comparison Group Design. The author's preference is to attempt to obtain a comparison group from surrounding schools or districts.

Sillito and Wilde (1983, p. 231-235) consider traditional program evaluation designs and identify some limitations. In addition to the above noted problem of obtaining comparison groups there are others:

> a) "Programs appropriate to gifted students invariably call for curricula which are qualitatively different from and represent a considerable extension of the regular curriculum. This renders it impossible to construct a single test instrument which is fair to both groups."
> b) "Even in a group of gifted students, the variability from student to student in kinds of gifts makes it unlikely that a test instrument can be devised which will lead to meaningful comparison between two experimental programs, with each group being the control for the other."
> c) "Another complicating factor . . . is . . . that at least some part of the student's curriculum ought to be individualized."

Another limitation, even in the weaker comparison group designs is that the control group will either differ very substantially from the gifted group in terms of the individuals who comprise the group, or does not function as a group and hence will have a considerable variability of treatment.

In terms of a goals-based model, whether for gifted individuals or a group of such individuals, it is difficult to know whether the goals are worthwhile and whether they may be attributed as an effect of the program. The product based model raises questions such as the relevance of the product to the goals, the value of the product, evaluation of process skills acquired, and whether the quality of product is attributable to the program.

Barnette (1983) noted the difference in philosophy and position between the hypothetico-deductive approach (experimental) and later naturalistic methods in the way that variability is viewed and processed. The experimental researcher regards any variance not specifically attributable to the treatment as error variance. This variance has the effect of concealing treatment effects so that significant results are found less often. It is a serious disadvantage,

> "Such as approach, although having a place in the arsenal of inquiry methods in education, has been criticized as being too narrow in scope and, thus lacking in maximization of useful information for evaluation of complex educational programs", (p. 26).

Naturalistic evaluation recognizes the great variability in the educational setting, attempts to document this variability, uses it to describe the complex set of variables and interactions and use this variability to interpret program events and their relation to outcomes. Without neglecting the useful quantitative information, naturalistic evaluation adds qualitative information and thus broadens the information base. Barnette claims that the naturalistic evaluation movement is neither a move away from quantitative to qualitative evaluation nor that qualitative evaluation is necessarily unreliable and non-objective.

Callahan (1983, p. 3) draws attention to evaluations which conclude that a program has/has not been effective, yet provide no information about the nature of the program and why it was or was not effective. She attributes the deficit to the fact that the program has not been adequately described. Callahan also draws attention to the demands for evaluation in the short term even though goals are quite likely to be long term.

## Improvements Suggested

A number of the researchers cited in the foregoing (and others) have offered solutions for the deficiencies noted in evaluating education programs for gifted students. In addition to these there are suggestions for improving educational evaluation generally which also has relevance to the problems encountered in evaluating programs and student outcomes in educating gifted students.

Renzulli's suggestion that programs have a sound legal, moral and philosophical base would remove some of the urgent pressures to evaluate prematurely. These are essentially the same bases on which regular education programs are supported. Thus the articulation with traditional education programs would likely receive increased attention, with some consequent improvement of both.

Weiss and Gallagher noted that much more formative evaluation is needed. Callahan implied the same, "Program developers must have the freedom to alter their plans as they receive further information . . ." because of the requirements of good program development. Renzulli (1975 p. 31) spoke of formative evaluation as, "an essentially simple concept [which] is emerging as one of the most powerful in present day thinking about evaluation." There is consensus that programs for gifted students require formative evaluation and that the need is greater because so many of the programs are in the developmental stages.

Callahan (1983) after concluding that the two basic approaches, experimental/control group, and behavioral objectives designs are not feasible for evaluating programs for gifted students, suggests a "Reasonable Person" approach based on the "Reasonable Man" concept in English Common Law. In essence the approach seeks information from enough independent sources so that a reasonable person may be convinced of the conclusions reached. A variety of instruments such as tests, rating scales, or several judges are used to substantiate program impact. If several of such sources attest to the same conclusion, it is more likely that reasonable people will agree. Of course consistent efforts should be made to ensure as well as can be, the validity and reliability of the evidence from each of the multiple sources. Multiple methodologies should be considered concurrently with multiple instruments. Among the alternatives are a time series design, using students as their own control or students from a matched school.

The time series design consists of dividing a group of students into a small number of subgroups, the curriculum into an equal number of components with the time assigned for each component being approximately the same. This organization provides the opportunity for the program impact in one subgroup to be compared with other subgroups not having received the treatment. For example, if a group of gifted students is divided into three (comparable) subgroups A, B, C; the curriculum into three components M, N, O; and the term into thirds I, II, III, the following organization could be established.

|  |  | Time Period |  |  |
|---|---|---|---|---|
|  |  | I | II | III |
| Subgroup | A | M a | N b | O c |
|  | B | N d | O e | M f |
|  | C | O g | M h | N i |

The comparison group would be:

a) Time period I: a with d and g (singly or combined)
d with a and g (singly or combined)
g with a and d (singly or combined)

b) Time period II: b with d
e with b
h with e

c) Time period III: none

Using students as their own controls is possible in certain situations. In the revolving door model, the productivity of students in the period when they are "in" the program should be compared with when they are not in the program. Student outcomes in a period of time prior to entry into a program could be compared with a subsequent period in the program.

Using students from a matching school is suggested as a "last resort".

The alternate methodologies suggested by Callahan are not offered as ideal procedures but as alternative sources of information to convince the "reasonable person". The technical term used for this multiple source of information is called "triangulation". Callahan's suggestion is triangulation of both instruments and methodologies.

Callahan further suggests that a comprehensive evaluation should enlarge the scope of questions asked and should provide a rich description of the program which will permit an understanding of the program components which produce the effects discovered.

Barnette (1983) argues for an in-depth description for the same reason and also so that another educator considering adoption of the program would have a basis for prejudging its suitability.

Barnette also advocates an extension of evaluation using a "naturalistic" approach which he describes as a cyclical information gathering process which provides for: a) the description, b) identification of "evaluation questions, issues and concerns", c) collection of preliminary general information, d) developing hypotheses from the questions, issues and concerns, e) verifying hypotheses and recycling.

"The intended result," he says,

> "is an evaluation which deals with the most important issues, strategies, outcomes of an educational program operating within the context of a dynamic social system, based on collection, organization, analysis, and interpretation of the

> most focused, reliable, and valid qualitative and quantitative information available."

He notes that, especially in the naturalistic approach to evaluation, it is important that evaluation is concurrent with program planning so that valuable direct observations are recorded and not trusted to faulty recollection.

Naturalistic evaluation focuses on interactions both human-human and human-material. The human-human interactions take place among: program administrators, program support staff, teachers, students, parents, resource persons, community, and evaluators. The materials listed are: physical facilities, library resources, permanent equipment, and disposable equipment. After citing Guba (1978), Barnette says,

> "Naturalistic evaluation methods tend to have a phenomenological base; focus on description and understanding; have as their purpose the discovery and verification of propositions; take a holistic view of the system or program being evaluated; work from an emergent, variable design mode; relate to the program in a selective rather than intervention manner; deal with multiple realities; and consider values an important set of variables to examine."

Barnette lists some naturalistic approaches:

a) "responsive evaluation (Stake, 1975)",

b) "illuminative evaluation (Parlett and Hamilton, 1976)",

c) "connoisseurship evaluation (Eisner, 1975)",

d) "judicial evaluation (Wolf, 1975)", and

e) "utilization focused evaluation (Patton, 1978)".

Barnette (op. cit.) uses the context, input, process, and product, (CIPP) approach of Stufflebeam, et. al. (1971) as a framework for the general kinds of questions evaluation would pose to gain the information required. In relation to context the questions center on student and teacher needs, external and internal support and available resources. The input questions relate to appropriate strategies, their feasibility, personnel and structures needed. Process questions focus on implementing the plan, the interactions occurring, modifications made and their effect. Product related questions center on achievement of objectives and the reasons why they were or were not attained, unanticipated outcomes and reasons therefore, the changes in interactions both historical and prospective if program is repeated or replicated.

Data collection should utilize the most appropriate range of methods to obtain the kind and scope of data required and should be unobtrusive. Macro-data, i.e. data about issues, concerns, and intents is gathered first to develop categories of data and to develop areas of more intensive investigation. Data may be obtained from:

- documents both internal and external to the program: program statements, proposals, interim reports, logs, budget expenditures.

- observation of the program. In this activity the thrust is to obtain in-depth data about functioning and interactions.
- interviews.
- surveys and other instrumental approaches.

The amount of data collected in an evaluation of this kind is massive and will require a simplifying system of categorization.

The role of evaluator is enhanced in naturalistic evaluation and the skill requirements are raised. The evaluator must have the ability to see organizations as an entity, to rely on tacit knowledge, to perceive what is expert and unique in a program, particularly in an innovative one.

Barnette concludes that,

> ". . . the naturalistic evaluation philosophy and methods have a great deal to offer in the evaluation of gifted and talented educational programs. Such programs tend to be complex, innovative, and dynamic. They are . . . influenced by . . . political forces . . . tend to be very visible . . . but not well understood by persons not directly involved in the program. Used appropriately, naturalistic evaluation can expect to provide not only a description of what happened in the program but more importantly, why things happened and what things need to be done to maintain or improve the program", (p. 36).

If the above notes abstracted from the literature on evaluating programs for gifted students appear to reflect a shift from evaluation as a science to evaluation as an art attitude, it is only partly because education for gifted students presents unique demands. Program evaluation development in education generally appears to be moving toward more emphasis on providing more detail on the program and its organic wholeness and less on the carefully crafted antisepsis of experimental design. Olivia Saracho (1982) addresses the topic, "New Dimensions In Evaluating The Worth Of A Program". She notes, "Although come evaluations are better than others, an exemplary model has not been developed." Her sampling of then current views includes: use of descriptive data and personal values to judge the worth of a program; persuasiveness of an evaluation is an indication of its validity; evaluation as a means of understanding the education system; as a basis for rewarding merit, or to improve curriculum balance among the rational, the intuitive, and the humane. She notes for general education evaluation purposes, the illuminative and responsive models. Evaluation reports "should be descriptive, declarative, holistic and rich". Her "Ideal Evaluation Approach" would combine an expert's best ideas with "information and input from the people involved in the program". Input data would include, "observations, dialogues, anecdotes, quotations, questionnaires, opinionnaires, interviews, slides, photographs, samples of student's work, student's logs, cassette tapes, videotapes, and films [and] . . . objective data". She concludes that the ideal model would be used to produce an evaluation which would explain what causes program effects.

Hamilton (1980) wrote about "Evaluating Your Own Program". "The place to begin," he wrote, "is with the informal means you already use to evaluate a program." He states three purposes, accountability, program improvement, and dissemination. The two basic questions are, Does it work? and How does it work? He makes a point of rich documentation and distinguishes it from evaluation. Documentation is the information base recorded in usable form; evaluation is the interpretive and judgemental processes usually based on documentation. Outcome evaluation, does it work?, and process evaluation, how does it work? may, according to Hamilton have much the same documentation but differ in timing and in the audience served.

Jones and Sherman (1980) report a comparison of two models in evaluating an English classroom. Eisner's Educational Criticism Model and Flander's Interaction Analysis. The report concludes,

> "In the final analysis, the most striking difference between these two approaches to the description of educational life is that one [Flander's] attends mainly to the incidence of behavior while the other [Eisner] attends to the social meaning of action", (p. 557).

Evaluation of Student Outcomes

Student outcomes, program impact, and program effectiveness are closely related, but not quite synonymous, terms having a common concern for assessing student growth. Program effectiveness refers to achievement of program goals which usually relate to students. Program impact refers to all student related outcomes, both anticipated and unanticipated. Student outcomes relate specifically to student goals and objectives some of which may be individual ones, particularly in programs for gifted students. In the evaluation literature surveyed the reference is, in the main, to program effectiveness or program impact both by intent and by terms used. Even so there is a dearth of attention.

Already noted is the lack of valid measures. Callahan notes the problem in the article, cited above, on Issues in Evaluating Programs for the Gifted. Archambault provides advice to those who must use out-of-level standardized tests, but his first item of advice is to avoid their use. Hamilton says,

> "The trade-off in using standardized tests is that they may not tap the special learning taking place in an innovative program. Their use is most safely limited to demonstrating that those involved in the program did not fall behind the conventional classes in their performance on

> standard measures of academic achievement", (p. 549).

Hamilton notes other data gathering devices, attitude measures, performance tests, and use of expert judges. Instruments and sources could include: a) questionnaires, b) interviews, c) group interviews, d) "unobtrusive measures" [by which he appears to mean observations such as improved attendance], e) journals or written reports, f) testimonials, g) other observations.

English (1980) proposes curriculum mapping as a means for identifying the actual curriculum,

> "The actual curriculum is the one the teacher employs in the classroom. That is the one the students encounter . . . the guide may be totally misleading . . . Mapping makes a simple but profound break with traditional procedures . . . mapping supplies an important ingredient that can make curriculum more effective", (p. 558).

Although Fenwick is writing about curriculum generally, it is easy to draw the inference for educating gifted students. For gifted students the curriculum is modified. Further modifications are made to accommodate subgroups of gifted students and still further adjustments provide for individualization. If evaluation hopes to relate student outcomes to treatment variables, the variation in the actual curriculum must be taken into account. Curriculum mapping which records curriculum content, emphasis and sequence is an important step in ensuring validity of assessment indicators.

A set of curriculum modifications for gifted students quite generally acknowledged relates to creativity. The provincial task force definition (1) includes it as do most local education agencies which have a program for gifted students. Torrance (1979) observes that there are complexities, "I recognized . . . that creative behavior requires more than creative thinking skills. Motivation and skills are also essential", (p. 11).

In addition to discussing the use of The Torrance Test of Creative Thinking (TTCT) and other tests of his own construction which measure various aspects of creative thinking, Torrance considers tests of motivations and creative skills. Measures of creative motivations discussed include Creative Motivation Scale (Torrance, 1971) and What Kind of Person Are You? (Khatena and Torrance, 1976). The measurement of creative skills has relied most successfully, he reports, on biographical inventories. He considers the following: Alpha Biographical Inventory (Taylor and Ellison, 1967), TTCT (Torrance, 1966), Thinking Creatively in Action and Movement (Torrance, E. Paul and Gibbs, M.S., 1977), and Something About Myself (Khatena and Torrance, 1976).

Aylesworth (1983) suggests guidelines for selecting instruments for use in evaluating programs for the gifted. He notes that there are many types of instruments which can be used as well as tests. These include questionnaires, rating scales, interview schedules, and their careful selection is just as important as the selection of tests. Validity is the most important criterion. It presupposes that specific evaluation questions have been asked. "Without these questions, one cannot determine the appropriateness of any instrument", (p. 39). In terms of procedure, "One must begin with the evaluation question and then determine if the instrument provides information to answer it." The three types of validity should be considered. These are construct, criterion, and content. The first two are technical questions, the last is user related. It, "involves the careful, detailed comparison of questions asked on the instrument being considered to the information sought", (p.40). The reliability, i.e. the precision of the instruments, is of importance.

> "Unfortunately, the issue of comparing reliability instruments does not arise frequently when selecting instruments for evaluating gifted programs. Most of the appropriate instruments have no reliability estimates to compare."

Reliability can be improved by asking the same question in a number of different ways and by asking the question of a larger number of people. It is noted, however, that in spite of these measures, the information will always be imprecise to an unknown extent. Finally, the information must be capable of interpretation, if it is to serve its purpose in answering the evaluation question.

Reis (1983) reports assessment of Enrichment Triad Model and Revolving Door Identification Model programs over a period of five years.

# BIBLIOGRAPHY

Alvino, James; McDonnel, Rebecca C. and Richert, Susanne "National Survey of Identification Practices in Gifted and Talented Education." Exceptional Children, Oct., 1981.

Archambault, Francis X. Jr. "Measurement and Evaluation Concerns in Evaluating Programs for the Gifted and Talented." Journal for the Education of the Gifted, Fall, 1983.

Aylesworth, Martin "Guidelines for Selecting Instruments in Evaluating Programs for the Gifted." Journal for the Education of the Gifted, 1983, 7:1.

Barnette, J. Jackson "Naturalistic Approaches to Gifted and Talented Program Evaluation." Journal for the Education of the Gifted, Fall, 1983.

Callahan, Carolyn M. "Issues in Evaluating Programs For The Gifted." Gifted Child Quarterly, Winter, 1983.

Clarke, S.C.T. and Nyberg, V.R. Programs for Gifted Pupils in Edmonton Catholic Elementary Schools. Alberta Education, December, 1984.

English, Fenwick, W. "Curriculum Mapping." Educational Leadership, April, 1980.

Hamilton, Stephen F. "Evaluating Your Own Program." Educational Leadership, April, 1980.

Jones, Keith and Sherman, Anne "Two Approaches to Evaluation." Educational Leadership, April, 1980.

Khatena, Joseph and Torrance, E. Paul Manual for Khatena - Torrance Creative Perception Inventory. Stoelting, Chicago, 1976.

Renzulli, Joseph S. A Guidebook For Evaluating Programs For The Gifted And Talented. Ventura County, Ventura, California, 1975.

Renzulli, Joseph "Thinking Through the Evaluation Process: an Interview with Dr. Joseph Renzulli," by Thomas M. Buescher, (Assoc. Ed.). Journal for the Education of the Gifted, Fall, 1983.

Saracho, Olivia "New Dimensions In Evaluating The Worth Of A Program." Education, 1982, 103:1.

Sillito, M. T. and Wilde, W. D. Educating the Gifted. Planning Services, Alberta Education, Edmonton, Alberta, 1983.

Taylor, C.W. and Ellison, R. L. Manual for Alpha Biographical Inventory. Institute for Behavioral Research in Creativity, Salt Lake City, 1967.

Torrance, E. Paul The Torrance Test of Creative Thinking: Norms Technical Manual. Research ed. Lexington, Mass., Personnel Press, 1966.

Torrance, E. Paul Technical-Norms Manual for the Creative Motivation Scale. Georgia Studies of Creative Behavior, University of Georgia, Athens, 1971.

Torrance, E. Paul and Gibbs, M. S. Technical-Norms Manual for Thinking Creatively with Action and Movement. Georgia Studies of Creative Behavior, University of Georgia, Athens, 1977.

Torrance, E. Paul Issues in Gifted Education. National/State Leadership Training Institute on the Gifted and Talented, Ventura County, Ventura, California, 1979. (Brief #8)

Weiss, Patricia and Gallagher, James Report on Education of Gifted, Volume II, Program Effectiveness, Education of Gifted and Talented Students: A Review. Frank Porter Graham Child Development Center, University of North Carolina at Chapel Hill, North Carolina, 27514, 1983.

# APPENDIX B

Strathcona County #20

# CHALLENGE I.E.P.

## Academic Summary

Name____________________

Grade ______________ Age______

Challenge Center________________

School Year__________________

Home Room Teacher_______________

Home Phone Number_______________

RECORDS

Ability Tests

| Name of Test | Date | Coded Score |
|---|---|---|
| | | |
| | | |
| | | |

End of Year Grades

| | | | | | |
|---|---|---|---|---|---|
| Language Arts | | | Science | | |
| Math | | | Social Studies | | |

Teacher Ratings in Selection

| Scale | Score | Comments |
|---|---|---|
| Learning | | |
| Motivation | | |
| Creativity | | |
| Leadership | | |

Relevant Comments from Selections

____________________

____________________

____________________

____________________

## Interests

Indicate the students general areas of interest. Where possible, indicate particular topics, issues or areas of study in which the student would like to do advanced work. Sources of this information should be noted, i.e. student interview, interest inventories, etc.

## SUMMARY and RECOMMENDATIONS (goals)

Indicate how the students individual needs/strengths will be addressed within the year plan for the Challenge Program - include social growth, work/study skills and academic/talent areas.

NOTE: This form is intended to serve an an ongoing programing guide. Please add to it as required

Strathcona County #20

## Challenge Class - Observations/Specific Plans Worksheet

| Communication | Not Observed | Seldom Observed | Moderate | Strong | Outstanding |
|---|---|---|---|---|---|
| Is able to follow the progress of class discussion | | | | | |
| Is understood in class discussion | | | | | |
| Brings varied and/or interesting facts or insights to class discussion | | | | | |
| Reads a wide range of materials for information and pleasure | | | | | |
| Communicates precisely and accurately | | | | | |
| Demonstrates richness of expression fluency, use of analogy, metaphor and so on | | | | | |

| Research and Organization | Not Observed | Seldom Observed | Moderate | Strong | Outstanding |
|---|---|---|---|---|---|
| Curious and interested in obtaining all relevant facts and/or arriving at the truth of the matter | | | | | |
| Realizes when more information is necessary | | | | | |
| Generates penetrating and formative questions | | | | | |
| A keen and alert observer; effectively extracts information from sources | | | | | |
| Employs a variety of resources in research | | | | | |
| Is able to arrive at own conclusions | | | | | |
| Keeps a well-preserved and orderly binder | | | | | |
| Completes tasks as assigned or agreed | | | | | |
| Brings resources/materials to class as assigned or agreed | | | | | |
| Uses classtime efficiently | | | | | |
| Plans or maps out the foreseen steps of a project | | | | | |

This form is intended to provide an ongoing analysis of student progress during the year and should complement the ancedoted records.
Suggested Code (Term 1-pencil Term 2-blue Term 3-red)

| Current Year Progress | | | | | |
|---|---|---|---|---|---|
| Language Arts | | | Math | | |
| Social Studies | | | Science | | |

Communication with Homeroom Teacher and Parent(s) Regarding Student Progress

| Date | Contact |
|---|---|
| | |

Specific Plans and Recommendations

| Date | Plans | Goals Attained |
|---|---|---|
| | | |

| Motivation | Not Observed | Seldom Observed | Moderate | Strong | Outstanding |
|---|---|---|---|---|---|
| Curious: shows desire for further knowledge or understanding of things, people, events, etc. | | | | | |
| Shows interest and enthusiasm over new ideas | | | | | |
| Follows through with work after initial excitement has faded, without large amounts of external motivation | | | | | |
| Becomes intensely absorbed in activity | | | | | |
| Looks at work as potential for improvement rather than burden or drudgery | | | | | |
| Shows independence of action and self-initiation of learning | | | | | |
| Sets high standards for self | | | | | |
| Seeks out challenge, often preferring complexity to simplicity | | | | | |
| Motivated by self-actualization needs | | | | | |

| Critical thinking | Not Observed | Seldom Observed | Moderate | Strong | Outstanding |
|---|---|---|---|---|---|
| Recognizes the importance of a knowledge base in discussion | | | | | |
| Integrates varieties of knowledge | | | | | |
| Rapidly grasps underlying principles, theories or generalizations | | | | | |
| Distinguishes fact from opinion or what might be true from what must be true | | | | | |
| Effectively uses deductive and inductive thinking skills | | | | | |
| Analyses complicated material | | | | | |
| Expresses and accepts constructive criticism | | | | | |
| Has a high tolerance for ambiguity or lack of closure, but nonetheless questions things that don't make sense | | | | | |
| Develops specific criteria for self-evaluation | | | | | |
| Shows pleasure in this intellectual activity | | | | | |

| Social and Affective | Not Observed | Seldom Observed | Moderate | Strong | Outstanding |
|---|---|---|---|---|---|
| Participates effectively in group discussion and decisions | | | | | |
| Adheres to or productively helps to change classroom procedures | | | | | |
| Works effectively in a group | | | | | |
| Willing to share work, insights or information which may be helpful to other class members | | | | | |
| Shows a sensitivity to others' feelings | | | | | |
| Concerned with values, ideals and justice | | | | | |
| Carries responsibility well | | | | | |
| Shows willingness to explore self | | | | | |
| Flexible and open in manner | | | | | |

| Creativity | Not Observed | Seldom Observed | Moderate | Strong | Outstanding |
|---|---|---|---|---|---|
| Exercises skill in fluency, flexibility originality and elaboration | | | | | |
| Playfully manipulates ideas | | | | | |
| Shows humour and imagination | | | | | |
| Takes risks | | | | | |
| Is willing to consider new ideas | | | | | |
| Is able to get good intellectual 'mileage' from limited resources | | | | | |
| Shows pleasure in this intellectual activity | | | | | |

# APPENDIX C

# APPENDIX C

## Further Reading

The decision, long considered and deliberately made, to reduce references in the guidebook to an absolute minimum, is based on two considerations. First is that use of the guidebook as a practical aid to evaluation in a specialized educational effort should be impeded as little as possible. Second, because the guidebook is founded more on current practice and specific field development and tryout experience than on the literature, even careful attention to the latter would mislead because the field experience roots of the guidebook would go unnoticed. The decision to reference minimally, although advantageous to the intended use of the guidebook, does not address the legitimate needs of some readers for further exploration.

This appendix is an effort to attend to the need some readers may feel for direction in exploring the fascinating field of education for the gifted. It is intended to provide references to themes rather than more specific points of interest and it is confined to publications which Alberts Education has, or is, making available to all schools. First is the literature survey developed as one of the three foundations for this guidebook and included as appendix A. Second is Educating The Gifted, published in 1983 abd supplied to schools by Alberta Education. Third is Educating Our Gifted and Talented Students in Alberta: A Resource Manual, published in 1985 and being distributed by Alberta Education. These three publications are well referenced and should be readily accessible in each school in the province and should prove helpful. Notwithstanding this, the serious/curious reader is encouraged to explore even more widely in current periodicals and recent publications which abound.

The following chart displays further reading references to some of the major themes encountered in the guidebook.

## APPENDIX C

| GUIDEBOOK THEMES | Appendix A | Further Reading | |
|---|---|---|---|
| | | Educating the Gifted | Resource Manual |
| Chapter II: The complex and ideosyncratic nature of giftedness | pp. 4, 5, 8 | Chapter II (all); pp. 10-58, 90-91, 151 | pp. vi, 1-9* to 1-13, 1-17 to 1-24 |
| Evaluation problems stemming from the nature of giftedness | pp. 1-27; specifically pp. 3, 4, 5, 6, 8, 21 | Chapter X pp. 230-238 | Chapter Five (all) specifically pp. 5-1 to 5-12 |
| Nature and use of presage* data in programming and evaluation | Not explicit | Not explicit but see: models of giftedness pp. 44-50; Chapter III (all) pp. 59-86; Chapters VI, VII | Not explicit but see: Chapter Four (all) pp. 4-1 to 4-46 |
| Chapter III: Special problems on program evaluation | pp. 1-27 | Chapter X pp. 230-238 | |
| Formative or summative evaluation | pp. 7, 11, 12, 14, 17 | pp. 225-226 | pp. 5-4, 5-10 |
| Evaluation procedures: suggested strategy | pp. 13-27<br>specifically pp. 24, 27 | | pp. 5-2 to 5-9 |
| Chapter IV: Special problems in evaluating student related outcomes | pp. 1-27<br>specifically pp. 22-23 | pp. 221-230 | |
| Student-related-outcomes evaluation procedures | p. 25 | | p. 5-3 |
| Chapter V: Implications | pp. 1-27 as a good source from which implications may be drawn | See implications section in Chapters: II pp. 57-58, IV pp. 117-118, V p. 148, VI p. 163. VII p. 176, VIII pp. 196-197, IX pp. 217-219, X p. 250 | |

---

* This method of numbering pages in the Resource Manual was used in the final pre-press draft so that if page numbering is changed during printing, the reference should still be useful. The 1-9, for example, refers to page 9 of Chapter I.

** Presage data is in essence "current wisdom" of the knowledgeable on any particular matter. Its use is so common in matters such as curriculum and evaluation as to go not only unnamed but unnoticed as well. Increased reliance on it in programs for the gifted suggests more explicit treatment is warranted in that setting in which practically every decision is based largely upon it.

# APPENDIX D

# Figure III - 3
Example
of
Planning Worksheet
(Calgary Evaluation)

| Evaluation question | Reported to | Purposes to be served | Data rquired | Sources of data | Collection of data | Assessment by | Reporting comments |
|---|---|---|---|---|---|---|---|
| Definition:<br>1. Is system's definition congruent with:<br>a. Alberta Ed.'s definition | CBE Administration | Identification policy program review | Description of operating definition<br>Alberta Ed.'s definition | Documents from EAS G coordinator, staff, CBE Administration; Principals Teachers of Gifted Students; Alberta Ed. Publications | Interview protocol<br>" "<br>" "<br>" " | External evaluators/ others, panel including system admin., eval., school admin., teachers of gifted students | |
| b. the literature | CBE Administration | Policy/program review | Definitions from literature | Educating the Gifted<br>Sillito&Wilde<br>Alb. Ed. Publication<br>other literature | Review literature | External evaluators | |
| c. what CBE is putting into place? i.e. with current/ proposed identification procedures | CBE Administration | Procedures review/ program review | Current/proposed identification procedures and structure | Documentation: EASG coordinator, staff, psychologists, school admin. school resource group | Interview protocol<br>" "<br>" "<br>" " | External evaluators with input from panel of CBE admin., school admin., teachers of gifted students | |